

curso de matemáticas

J. LELONG-FERRAND , J. M. ARNAUDIÈS

TOMO I

álgebra

editorial reverté, s.a.

Hermann Langguth

Curso de matemáticas

Tomo 1

ÁLGEBRA

Curso de matemáticas

Tomo 1

ÁLGEBRA

Jacqueline LELONG-FERRAND

Professeur à l'Université de Paris VI

Jean-Marie ARNAUDIÈS

Professeur de Mathématiques Spéciales au Lycée Kléber à Strasbourg



EDITORIAL REVERTÉ, S. A.

Barcelona - Bogotá - Buenos Aires - Caracas - México - Rio de Janeiro

Título de la obra original

Cours de Mathématiques — Tome 1
ALGÈBRE, 2^e édition revue et corrigée

Edición original en lengua francesa publicada por
DUNOD, Paris-Bruxelles-Montréal

Copyright © DUNOD, 1^{re} édition

Copyright © BORDAS, 2^e édition

Versión española por el

Dr. José Pla Carrera

Doctor en Matemáticas, Profesor de la Facultad de Matemáticas
de la Universidad de Barcelona

Revisada por el

Dr. Enrique Linés Escardó

Catedrático de la Facultad de Ciencias de la Universidad de Madrid

Propiedad de EDITORIAL REVERTÉ, S. A. — Encarnación, 86 — Barcelona(24)

Reservados todos los derechos. Ninguna parte del material cubierto por este título de propiedad literaria puede ser reproducida, almacenada en un sistema de informática o transmitida de cualquier forma o por cualquier medio electrónico, mecánico, fotocopia, grabación u otros métodos sin el previo y expreso permiso por escrito del editor.

Edición en español

© EDITORIAL REVERTÉ, S. A., 1979

Printed in Spain — Impreso en España

ISBN - 84 - 291 - 5065 - X obra completa

ISBN - 84 - 291 - 5066 - 8 tomo 1

Dep. Leg. B. 20753 - 1979

Litoclub, S. A. - Nápoles, 300 - Barcelona-25

Advertencia

*Hemos querido que esta obra sea a la vez un manual práctico, que presente una exposición clara y detallada del programa, y un instrumento de reflexión para los estudiantes deseosos de profundizar ciertos aspectos de la teoría, y de iniciarse en problemas más elevados. A este fin, los desarrollos que pueden dejarse de lado en una primera lectura, o en una revisión, se hallan impresos en letra pequeña o indicados por un *. El texto no pretende, sin embargo, substituir al profesor, pues es a éste a quien incumbe el papel esencial de guiar al estudiante en el estudio del curso, y de indicarle los resultados importantes con miras a las aplicaciones: esperamos que esta obra podrá ayudarle en su tarea.*

Hemos procurado, en la medida de lo posible, que los capítulos fuesen independientes. Ello nos ha conducido a demostrar varias veces un mismo teorema, o a repetir, en ciertos lugares, el significado de definiciones que figuran en otros capítulos. Por otro lado, la preocupación de conservar, en lo posible, el texto ordinario independiente del de letra pequeña nos ha conducido también a ciertas repeticiones, que se encontrarán principalmente en el capítulo XI, en relación con la reducción de Jordan de las matrices cuadradas.

Las generalidades acerca de la teoría de conjuntos se han reducido al mínimo, y para un estudio más completo, remitimos a [4]. En los capítulos II y III, por el contrario, hemos creído necesario desarrollar ampliamente las bases indispensables en lo sucesivo.

En la parte de «Álgebra lineal» nos hemos limitado voluntariamente a los espacios vectoriales de dimensión finita, contentándonos con indicar, de paso, extensiones fáciles. Pues pensamos que — para el algebrista — las propiedades de dimensión finita son profundas. Y no hemos hallado interesante desarrollar ampliamente las consecuencias del axioma de la elección, ya que a este nivel el lector no puede deducir aplicaciones concretas. Pensamos que es en Análisis, a través de los espacios funcionales, que el estudiante se iniciará de forma atractiva en los espacios de dimensión

infinita (en un marco en que, por el contrario, el papel de una base algebraica no tiene la misma importancia).

En los capítulos V, VI y XIV hemos desarrollado y precisado, a veces, temas del programa que corresponden a la parte más tradicional del Álgebra. A pesar del abandono a que suele hallarse sometida, no la hemos sacrificado, aunque sólo sea para desarrollar en el lector la práctica de algoritmos útiles y el dominio del cálculo algebraico.

El signo ●, a principios de párrafo, indica un convenio o hipótesis, implícitamente aceptado a lo largo del párrafo. En el curso de un desarrollo puede indicar también un punto esencial para la comprensión del texto.

Damos sinceras gracias a Ediciones Dunod que, a pesar de las dificultades del texto, han realizado una tipografía casi perfecta.

Finalmente, damos por adelantado las gracias a todos aquellos lectores que tengan a bien hacernos partícipes de sus observaciones.

Modo de empleo

1) *El texto se halla dividido en capítulos, designados por cifras romanas.*

Cada capítulo se divide en párrafos. En un párrafo se hallan numerados linealmente: por un lado, las definiciones, y por otro, los teoremas y proposiciones. Las proposiciones se hallan indicadas simplemente por su número de orden y su enunciado no debe constituir un esfuerzo de memoria, puesto que sólo constituyen etapas.

Ejemplo

DEFINICIÓN VI.1.1 = definición n.º 1, del § 1, del capítulo VI.

TEOREMA VI.1.1 = teorema n.º 1, del § 1, del capítulo VI.

VI.1.2 = proposición n.º 2, del § 1, del capítulo VI.

Los corolarios de los teoremas se hallan indicados únicamente por un número de orden, si hay varios; en caso contrario, no se han numerado. Así se hablará del «corolario 1, del teorema VII.2.2» o bien del «corolario del teorema VIII. 3.1».

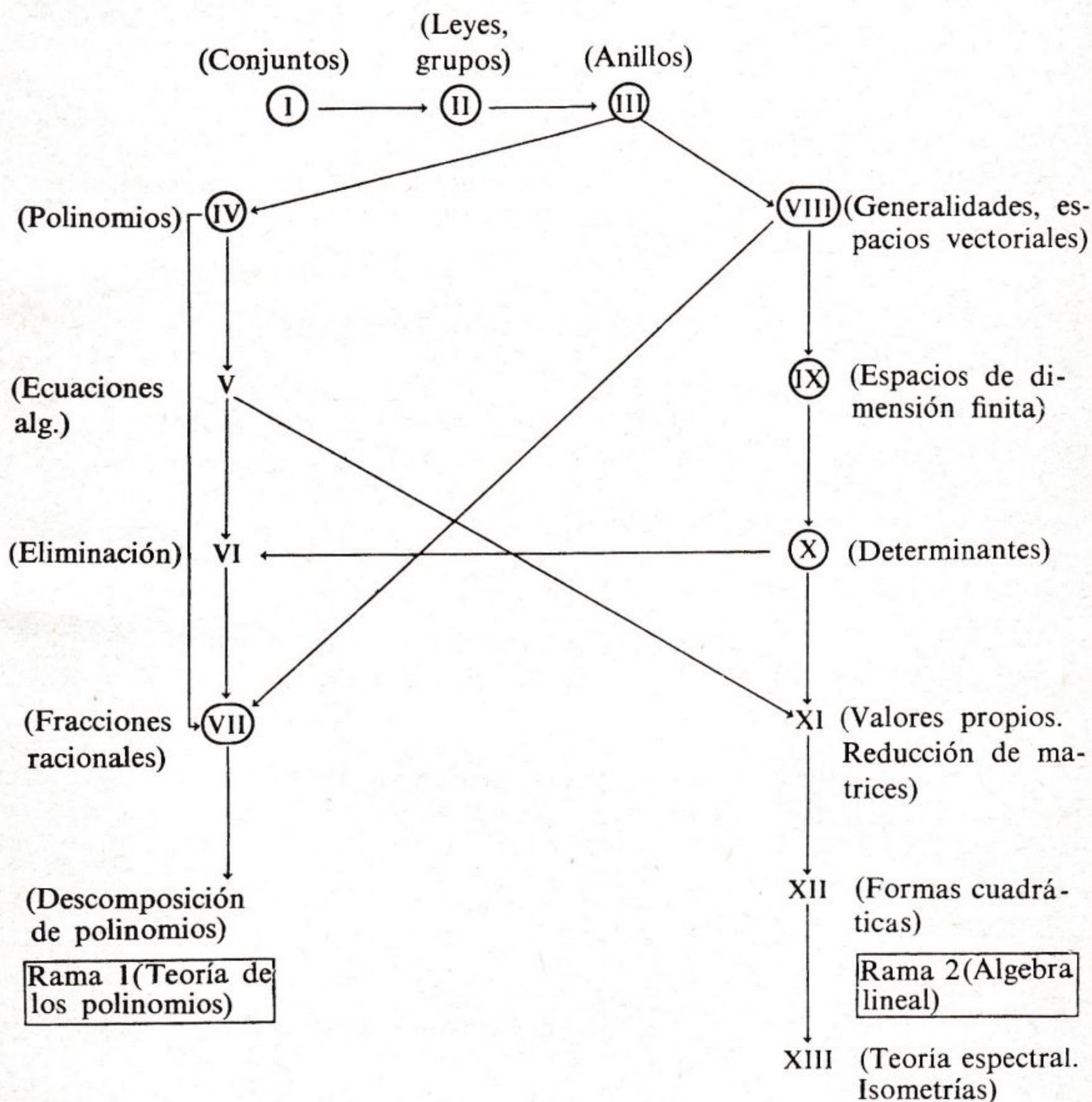
2) *La abreviatura c.q.d. indica el final de una demostración o de un razonamiento. La hemos utilizado a veces al final de un razonamiento para indicar de forma clara el cambio de tema.*

3) *Si el texto, en un cierto lugar, comporta una referencia a resultados demostrados ulteriormente, ello implica que dichos resultados son independientes de la cuestión tratada.*

4) *Nos hemos permitido dar ejemplos que utilizan nociones que no se dan en esta obra, principalmente nociones de Análisis o de Geometría.*

5) Al final del libro presentamos una recopilación de ejercicios, numerados según el capítulo al que pertenecen, y una recopilación de problemas de síntesis. Los ejercicios más difíciles se han indicado con el signo *.

Interdependencia de los capítulos



Hemos limitado con el signo ○ los números de los capítulos que deben asimilarse en un año de primer ciclo o de Matemáticas superiores.

Indice analítico

CAPÍTULO I. Vocabulario de la teoría de conjuntos	1
1 Nociones sobre formalización	1
2 Reglas de lógica formal	3
3 Cuantificadores	6
4 Operaciones sobre los conjuntos	9
5 Correspondencias y aplicaciones	13
6 Familias	20
7 Productos	23
8 Relaciones de equivalencia y conjunto cociente	26
9 Relaciones de orden	30
10 Enumeración	36
11 Conjuntos de base	45
CAPÍTULO II. Leyes de composición. Grupos	47
1 Generalidades	47
2 Propiedades de una ley de compensación	52
3 Axiomas de la estructura de grupo. Ejemplos de grupos. Homomorfismos	57
4 Subgrupo. Grupo engendrado. Grupo producto	61
5 Grupo cociente en el caso abeliano	66
6 Grupos cualesquiera: clases, subgrupos normales. Cociente	72
7 Grupos finitos. Grupo simétrico	76
8 Grupo que opera sobre un conjunto	81
CAPÍTULO III. Estructuras algebraicas en las que intervienen varias leyes	89
1 Generalidades	89
2 Generalidades sobre los anillos	91
3 Subanillos y anillos productos	100
4 Homomorfismos, ideales y anillos cociente	102
5 Divisibilidad de un anillo	109
6 Cuerpos	113
7 Cálculo en el cuerpo de los números complejos	120
8 Estructura de módulo sobre un anillo	127
9 Estructura de álgebra sobre un anillo conmutativo unífero	135

CAPÍTULO IV. Polinomios con una o varias variables	137
1 Definición de $A[X]$, propiedades generales	137
2 División euclídea. Propiedades aritméticas de $K[X]$ cuando K es un cuerpo conmutativo	143
3 Algoritmo de Euclides	149
4 Polinomios irreducibles (sobre un cuerpo)	154
5 Función polinomio. Raíces. Fórmula de Taylor	156
6 Relaciones entre los coeficientes y las raíces. Descomposición en $C[X]$ y $R[X]$	165
7 Nociones sobre $K[X_1, X_2, \dots, X_n]$	
CAPÍTULO V. Funciones simétricas. Ecuaciones algebraicas (teoría elemental)	183
1 Polinomios simétricos	183
2 Fórmulas de Newton	190
3 Ecuaciones de segundo y tercer grado	197
4 Ecuación de cuarto grado	204
5 Ecuación de grado mayor que 5	208
CAPÍTULO VI. Eliminación	215
1 Método de Cayley. Resultante de dos polinomios con una sola variable	216
2 Algunos ejemplos de cálculo de resultantes	223
3 Aplicación de la eliminación a la transformación de Tschirnhaus	230
4 Discriminante de un polinomio	237
5 Expresión de las raíces comunes a dos ecuaciones cuando su resultante es nula	241
CAPÍTULO VII. Fracciones racionales	249
1 El cuerpo $K[X]$	249
3 Descomposición en elementos simples sobre un cuerpo cualquiera K	253
3 Cálculo de las partes polares relativas a los factores de la forma $(X - a)^a$	257
4 Nociones acerca de las series formales	260
5 Ejemplos de cálculos prácticos	266
6 Integración de fracciones racionales	273
CAPÍTULO VIII. Espacios vectoriales	281
1 Generalidades	281
3 Caracterización de las bases de un espacio vectorial	287
3 Teorema de la dimensión finita	290
4 Espacios vectoriales y aplicaciones lineales: rango de una aplicación lineal	297
5 Dualidad	300
6 Lenguaje de la geometría afín	309
CAPÍTULO IX. Matrices	315
1 Matrices	315
2 Matrices y aplicaciones lineales	326
3 Cambio de base	336
CAPÍTULO X. Los determinantes y sus aplicaciones	343
1 Aplicaciones y formas multilineales	351
2 Determinantes	351
3 Ejemplos de cálculo de determinantes	361
4 Aplicación de los determinantes al estudio del rango de una matriz	368

5 Ecuaciones lineales	372
CAPÍTULO XI. Reducción de las matrices cuadradas y aplicación.....	383
1 Valores propios, polinomio característico	383
2 Subespacios propios	387
3 Polinomios de endomorfismos. Teorema de Hamilton-Cayley	394
4 Subespacios característicos	400
5 Endomorfismos diagonalizables	407
6 Endomorfismos nilpotentes. Factores invariantes. Reducción de Jordan	408
CAPÍTULO XII. Formas bilineales y formas cuadráticas	415
1 Generalidades acerca de las formas bilineales	415
2 Formas bilineales simétricas y formas cuadráticas	418
3 Ortogonalidad	423
4 Problema de la clasificación. Solución cuando $K = \mathbb{C}$ o $K = \mathbb{R}$	427
5 Espacio euclídeo.....	433
6 Proyecciones y simetrías.....	443
7 Grupo ortogonal, el grupo ortogonal real	445
CAPÍTULO XIII. Formas hermíticas. Teoría espectral. Isomerías de \mathbb{R}^n	455
1 Generalidades	455
2 Clasificación de las formas hermíticas sobre un espacio de dimensión finita....	459
3 Espacios prehilbertianos de dimensión finita	460
4 Proyecciones y simetría	465
5 Grupo unitario	466
6 Teoría espectral (formas hermíticas)	471
8 Teoría espectral (formas cuadráticas sobre un cuerpo cualquiera)	482
2 Isometrías de E_n (espacio euclídeo de dimensión n)	489
10 Isometrías vectoriales	492
11 Isometrías afines: teorema de prolongación	499
CAPÍTULO XIV. Polinomios de varias variables y aplicaciones geométricas	503
1 Anillos factoriales	503
2 Factorialidad de los anillos de polinomios	507
3 Correspondencias algebraicas. Homografías (en característica 0)	511
4 Hipersuperficiales algebraicas en \mathbb{C}^n ($n \geq 2$)	515
5 Curvas algebraicas y curvas algebraicas unicursales en \mathbb{C}^2	522
EJERCICIOS	527
PROBLEMAS	571
BIBLIOGRAFÍA	589
SÍMBOLOS UTILIZADOS EN ESTE TRATADO	591
ÍNDICE ALFABÉTICO	595

Curso de matemáticas

Tomo 1

ÁLGEBRA

Capítulo I

Vocabulario de la teoría de conjuntos

§ I.1 NOCIONES SOBRE FORMALIZACIÓN

La teoría intuitiva de conjuntos, tal como se había desarrollado a principios de siglo, conducía a *paradojas* molestas; son bien conocidas, por ejemplo, las paradojas debidas a la consideración del «conjunto de los conjuntos», o la que se refería: «al menor número entero que no se puede definir con menos de dieciséis palabras castellanas». Estas paradojas se deben al uso incontrolado del lenguaje corriente para construir nuevos objetos matemáticos hasta el infinito, y razonar sobre ellos, lo que introduce, por una parte, una excesiva riqueza de lenguaje y, por otra, todas las dificultades de la «lógica trascendental».

Lenguajes formalizados

Para mitigar estos inconvenientes se han inventado los *lenguajes formalizados* y las *lógicas formales*, que permiten construir la teoría de conjuntos mediante el *método axiomático*. De forma muy esquemática, una construcción formalizada y axiomática de dicha teoría se ajusta al siguiente modelo: se da un número reducido de signos lógicos y un número reducido de reglas que permitan, con la ayuda de estos signos y de las letras de los distintos alfabetos, escribir las «palabras permitidas». (A las palabras permitidas se les llama frecuentemente *agrupaciones*.) Se da un método que permite distinguir dos tipos de palabras permitidas: unas, llamadas *términos*, serán los representantes abstractos de los objetos acerca de los que se razonará, y las otras, llamadas *relaciones*, representarán las afirmaciones que se pueden hacer acerca de dichos objetos. Después se dan reglas que rijan el uso de

las relaciones y que permiten construir nuevas relaciones a partir de relaciones dadas, etc.; estas reglas son las reglas de la *lógica formal*. Hecho esto, la noción de verdad matemática se «relativiza» de la manera siguiente: se toma, como verdaderas *a priori*, un reducido número de relaciones, llamadas *axiomas*. A continuación se define la noción de *demostración*, que es un texto formalizado escrito siguiendo las reglas, que contiene términos y relaciones, y parte de uno o varios axiomas. Entonces se llama *verdadera* una relación si se la puede insertar en una demostración. Se observa, pues, que esta noción de verdad se reduce al acuerdo del espíritu consigo mismo y que, en principio, poco importa el significado «absoluto» de los axiomas.

Cuando se dispone de un lenguaje formalizado coherente para fundamentar una teoría (la teoría de conjuntos o cualquier otra teoría matemática), el desarrollo de esta teoría consiste en hallar las relaciones verdaderas, a las que se da el nombre de *teoremas*, *proposiciones*, *lemas*, *escolios*, etc.

Interpretación de los textos formalizados

Así como el trabajo de un ordenador no se reduce a la producción de «fichas perforadas artísticamente», tampoco las matemáticas se reducen al juego estéril de inventar un lenguaje formalizado coherente, para después escribir, al azar, las relaciones verdaderas de este lenguaje. Volviendo a nuestra comparación con el ordenador, el papel más importante de un matemático es una actividad de «software», es decir, de interpretación, orientación, y utilización de resultados. En efecto, los matemáticos —como los físicos y los administradores— tienen problemas propios, de aspecto muy concreto, y algunos permanecen sin resolver desde hace siglos. La formalización es sólo un instrumento necesario, según se ha comprobado, y cuyo uso queda justificado sólo por el éxito: éxito en la resolución de ciertos problemas técnicos, y sobre todo, en la transmisión de los conocimientos matemáticos, gracias a la claridad del texto, a su economía y a su universalidad. Por otro lado, la formalización es un instrumento muy flexible: hoy en día se considera que el estudio de una cuestión concreta está muy avanzado si se ha logrado plantearla en lenguaje formalizado (cf. Aplicación de las matemáticas a la administración, a los transportes ferroviarios, a las telecomunicaciones, a la programación de los viajes cósmicos, etc.).

No se debe, pues, olvidar que un lenguaje formalizado «útil» se halla *orientado de antemano*, y escrito para construir un *modelo abstracto* del problema que se pretende estudiar; por consiguiente, ni los axiomas ni los símbolos lógicos son arbitrarios ni gratuitos. Por este motivo, sólo un lector familiarizado con problemas propiamente matemáticos puede abordar con fruto el estudio de un lenguaje formalizado (estudio que no se realizará en esta obra.)

Aspecto práctico de la formalización

Antes de tratar de nuestro vocabulario, señalemos que es posible construir todas las matemáticas conocidas, hoy por hoy, con la ayuda de los axiomas y del lenguaje formalizado de la teoría de conjuntos. El axioma fundamental ⁽¹⁾ lo constituye la existencia de un conjunto, por lo menos, de naturaleza matemática: el de los números enteros. Gracias a este axioma, es posible aplicar los resultados abstractos de la teoría de conjuntos a cosas distintas del «conjunto de los árboles del "quartier latin"» cuya existencia plantea actualmente serios problemas, o del «conjunto de los reyes de Francia» (que no existe, ya que la relación de pertenencia no está definida para el elemento Luis XVII).

Pero en la práctica, es imposible escribir todas las matemáticas en lenguaje formalizado, ya que el más pequeño de los teoremas fáciles necesitaría libros enteros. Nos vemos, pues, obligados a utilizar abreviaturas y el lenguaje corriente, y nos contentamos con escribir textos «de los que estamos seguros» que se *podrían* formalizar.

En lo que sigue, nos limitaremos a enunciar las reglas de empleo de las relaciones y de los conjuntos (sin pretender fundamentar rigurosamente la teoría). Se puede consultar [4], capítulos I y II.

§ 1.2 REGLAS DE LA LÓGICA FORMAL

— Según hemos dicho en el § 1, una relación es *verdadera* si se la puede insertar en un texto demostrativo. Dada una relación A , se define su *contraria*, designada por $\text{no } A$; a $(\text{no } A)$ también se le llama *negación* de A . Por definición, A es falsa si $(\text{no } A)$ es verdadera.

Si existe una A tal que A y $(\text{no } A)$ son verdaderas, a la teoría se le llama *contradictoria*, y se demuestra que, entonces, toda relación de la teoría es verdadera. A pesar de que no ha sido demostrado, se admite generalmente que *la teoría de conjuntos es no contradictoria*, de modo que, para toda relación A de esta teoría, a lo sumo una de las relaciones A y $(\text{no } A)$ es verdadera.

No debe creerse que una de las relaciones A y $(\text{no } A)$ es forzosamente verdadera: podrían existir relaciones contrarias A y $(\text{no } A)$ tales que no fuese posible insertar ninguna de ellas en un texto demostrativo. Se dice entonces que A es *indecidible*. En la práctica, no encontraremos relaciones indecidibles.

— Dadas dos relaciones A y B , se define la *disjunción* de A y B , designada por: $(A \text{ o } B)$. Si una, por lo menos, de las relaciones A , B es verdadera, $(A \text{ o } B)$ es verdadera.

A la relación $((\text{no } A) \text{ o } B)$ se le llama *implicación de B por A* , y se designa por:

$$A \Rightarrow B.$$

⁽¹⁾ Fundamental, por lo menos, para los matemáticos.

Si A y $(A \Rightarrow B)$ son verdaderas, B es verdadera; si B es verdadera, $(A \Rightarrow B)$ es verdadera para toda relación A . Además admitiremos las relaciones que siguen, que proporcionan reglas de razonamiento:

- 1) $(A \Rightarrow A)$. Esta relación es muy interesante, ya que expresa que (¡aunque A sea indecidible!) $(A$ o (no A)) es siempre verdadera.
- 2) $(\dot{A} \text{ o } A) \Rightarrow A$
- 3) $A \Rightarrow (A \text{ o } B)$
- 4) $(A \text{ o } B) \Rightarrow (B \text{ o } A)$
- 5) $[A \Rightarrow B] \Rightarrow [(C \text{ o } A) \Rightarrow (C \text{ o } B)]$
- 6) $[A \Rightarrow B] \Rightarrow [(B \Rightarrow C) \Rightarrow (A \Rightarrow C)]$
- 7) $A \Rightarrow (\text{no } (\text{no } A))$.
- 8) $[A \Rightarrow B] \Rightarrow [(\text{no } B) \Rightarrow (\text{no } A)]$.

De entre las reglas anteriores, las reglas números 1), 2), 3), 4), 5) son *axiomas* en la mayor parte de las lógicas formales usuales.

— Si A y B son relaciones, se define la *conjunción* de A y B , designada por $(A \text{ y } B)$, como la relación:

$$\text{no}[(\text{no } A) \text{ o } (\text{no } B)].$$

Se dice que A y B son *equivalentes* si $(A \Rightarrow B)$ y $(B \Rightarrow A)$ son verdaderas; se escribe entonces: $(A \Leftrightarrow B)$. A las relaciones: $(A \Rightarrow B)$ y $(B \Rightarrow A)$ se les llama *implicaciones recíprocas*.

Se demuestra que las recíprocas de 7) y 8) son verdaderas, de modo que se tienen las equivalencias:

- 9) $A \Leftrightarrow \text{no } (\text{no } A)$
- 10) $(A \Rightarrow B) \Leftrightarrow ((\text{no } B) \Rightarrow (\text{no } A))$.

A la implicación: $(\text{no } B) \Rightarrow (\text{no } A)$ se le llama *contrarrecíproca* de $(A \Rightarrow B)$. En ciertos casos es más fácil demostrar $(\text{no } B) \Rightarrow (\text{no } A)$ que $(A \Rightarrow B)$.

Veamos algunas propiedades en que intervienen la disjunción, la conjunción, la implicación y la equivalencia:

- 11) $[(A \Rightarrow B) \text{ y } (B \Rightarrow C)] \Rightarrow [A \Rightarrow C]$ (transitividad de la implicación)
- 12) $\text{no } (A \text{ y } B) \Leftrightarrow [(\text{no } A) \text{ o } (\text{no } B)]$
- 13) $\text{no } (A \text{ o } B) \Leftrightarrow [(\text{no } A) \text{ y } (\text{no } B)]$
- 14) $[(A \text{ o } B) \text{ o } C] \Leftrightarrow [A \text{ o } (B \text{ o } C)]$ (asociatividad de «y» y «o»)

- 15) $[(A \text{ y } B) \text{ y } C] \Leftrightarrow [A \text{ y } (B \text{ y } C)]$
 16) $[(A \text{ o } B) \text{ y } C] \Leftrightarrow [(A \text{ y } C) \text{ o } (B \text{ y } C)]$
 17) $[(A \text{ y } B) \text{ o } C] \Leftrightarrow [(A \text{ o } C) \text{ y } (B \text{ o } C)]$
 18) $[(A \text{ o } B) \text{ y } (A \Rightarrow C) \text{ y } (B \Rightarrow C)] \Rightarrow C.$

La propiedad (18) constituye el razonamiento conocido como *disjunción de casos*. Su interés radica en que, si $(A \text{ o } B)$ es verdadera, no se sabe nada acerca de la verdad de A , ni acerca de la de B . Por ejemplo, ya que $(A \text{ o } (\text{no } A))$ es siempre verdadera, la relación $[(A \Rightarrow C) \text{ y } (\text{no } A \Rightarrow C)] \Rightarrow C$ es siempre verdadera (¡incluso cuando A es indecidible!).

Ejemplo

Queremos demostrar el teorema «pequeño» de Fermat: «Para todo entero n y todo número primo p , se tiene $n^p - n \equiv 0 \pmod{p}$ ».

Demostración. (Disjunción de casos):

- Si $n \equiv 0 \pmod{p}$, es evidente.
- Si no, n es primo con p , y la congruencia propuesta es equivalente a $n^{p-1} - 1 \equiv 0 \pmod{p}$, o también, designando por $\chi: \mathbf{Z} \rightarrow \mathbf{Z}/p\mathbf{Z}$ el homomorfismo canónico, a $(\chi(n))^{p-1} = \chi(1)$. Ahora bien, esta última relación es verdadera, puesto que el grupo multiplicativo $(\mathbf{Z}/p\mathbf{Z})^* = (\mathbf{Z}/p\mathbf{Z}) \setminus \{0\}$ contiene $p - 1$ elementos (cf. Ejemplo 2, p. 116). \square

Demostración por reducción al absurdo

Es la aplicación de la regla siguiente (que admitiremos):

Sea A una relación; si existe una relación B (no necesariamente verdadera) tal que las relaciones: $(\text{no } A) \Rightarrow B$ y $(\text{no } A) \Rightarrow (\text{no } B)$ son verdaderas, entonces A es verdadera.

Ejemplo

Queremos demostrar el siguiente teorema (cf. § IV.2):

$$\left. \begin{array}{l} (A, B, Q, R, Q', R' \text{ son polinomios con coeficientes} \\ \text{en un cuerpo conmutativo } K \text{ tales que:} \\ B \neq 0, A = BQ + R = BQ' + R' \\ \text{y } \text{gr}(R) < \text{gr}(B) \text{ y } \text{gr}(R') < \text{gr}(B) \end{array} \right\} \Rightarrow (Q = Q' \text{ y } R = R').$$

Demostración. Supongamos verdadero el primer término de la implicación. Entonces se ha de probar que el segundo es verdadero. Razonemos por reducción al absurdo, y supongamos falso el segundo miembro. Entonces tendremos:

$$(Q \neq Q') \quad \text{o} \quad (R \neq R')$$

lo que teniendo en cuenta las relaciones $B(Q' - Q) = R - R'$ y $B \neq 0$, implica:

$$((Q \neq Q') \quad \text{y} \quad (R \neq R'))$$

de donde:

$$\text{gr } B(Q' - Q) \geq \text{gr } B \quad \text{y} \quad \text{gr } (R - R') < \text{gr } B,$$

que es absurdo. Luego $Q = Q'$ y $R = R'$. c.q.d.

Análisis de este razonamiento: La teoría τ considerada es la de los polinomios con coeficientes en un cuerpo conmutativo K . El teorema propuesto estará demostrado si se prueba que el teorema $T : (Q = Q' \text{ y } R = R')$ es verdadero en la teoría τ' obtenida añadiendo a la teoría de los polinomios los axiomas: « A, B, Q, R, Q', R' son polinomios, $B \neq 0$, $A = BQ + R$, $A = BQ' + R'$, $\text{gr } (R) < \text{gr } (B)$, $\text{gr } (R') < \text{gr } (B)$ » (a estos axiomas momentáneos se les llama «hipótesis auxiliares»).

Designemos por P el polinomio $B(Q' - Q)$, que es también el polinomio $R - R'$, y por U la relación: $\text{gr } (P) \geq \text{gr } (B)$. Se ha demostrado

$$(\text{no } T) \Rightarrow U \quad \text{y} \quad (\text{no } T) \Rightarrow (\text{no } U).$$

Según la regla de razonamiento por reducción al absurdo, se obtiene que T es verdadero en τ' , y, por lo tanto, que el teorema propuesto es verdadero en τ . En la práctica, nos contentamos con la redacción que precede a nuestro análisis.

Nota. En muchos casos, una demostración llamada «por reducción al absurdo» se puede convertir en la propiedad (10), y consiste, simplemente, en demostrar $(\text{no } B) \Rightarrow (\text{no } A)$, para demostrar $(A \Rightarrow B)$.

§ 1.3 CUANTIFICADORES

Sea A una relación en la cual puede intervenir el término x ; (intuitivamente, A es una relación que depende del objeto x (incluso cuando éste no figura explí-

citamente en ella)). Para poner de relieve a dicho término, se escribirá $A(x)$ en vez de A . Dicho esto, escribiremos

$$(1) \quad \exists x, A(x)$$

para expresar la relación «la relación $A(x)$ es verdadera para un objeto x , por lo menos». Esta definición es totalmente intuitiva, y sólo describiremos las reglas de uso del símbolo \exists , llamado *cuantificador existencial*.

Sea $A(x)$ una relación dependiente de x . Escribiremos:

$$\forall x, A(x)$$

para expresar la relación:

$$(2) \quad \text{no } (\exists x, \text{no } A(x)).$$

Esta relación significa que la propiedad $A(x)$ es verdadera para todos los objetos x .

Al símbolo \forall se le llama *cuantificador universal*.

Las fórmulas (1), (2) y las reglas de la lógica permiten utilizar mecánicamente los cuantificadores.

Ejemplos

$$1) (\forall x, (A(x) \text{ y } B(x))) \Leftrightarrow ((\forall x, A(x)) \text{ y } (\forall x, B(x))),$$

$$(\exists x, (A(x) \text{ o } B(x))) \Leftrightarrow ((\exists x, A(x)) \text{ o } (\exists x, B(x))).$$

Por el contrario, $(\forall x) (A(x) \text{ o } B(x))$ no es equivalente a $(\forall x, A(x) \text{ o } (\forall x, B(x)))$.

$$2) (\exists x, (\forall y, A(x, y))) \Rightarrow (\forall y, (\exists x, A(x, y))).$$

Por el contrario, no se tiene la implicación recíproca.

Nota. Negar una relación que contenga cuantificadores puede resultar delicado, pues puede ocurrir que algunos cuantificadores se hallen sobrentendidos. Por ejemplo, la negación de:

$$(1) \quad (\forall \varepsilon > 0, \exists \eta > 0, |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon)$$

es:

$$(2) \quad (\exists \varepsilon > 0, \forall \eta > 0, \exists x, |x - x_0| \leq \eta \text{ y } |f(x) - f(x_0)| > \varepsilon)$$

Con todo rigor, la fórmula (1) debería escribirse:

$$(\forall \varepsilon > 0, \exists \eta > 0, \forall x, |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon).$$

Por un abuso de lenguaje permitido, el símbolo « $\forall x$ » no figura explícitamente en la relación (1). Sin embargo, la omisión de « $\exists x$ » en la relación (2) constituiría una falta evidente.

Problema de la elección

Intuitivamente, el problema es el siguiente: ¿Es posible demostrar, en una teoría dada, un teorema de la forma: $(\exists x, \forall (x))$, sin *construir*, por medio de un procedimiento descriptivo, un objeto x para el cual la relación $A(x)$ sea efectivamente verdadera? Esta pregunta ha suscitado gran número de polémicas. Grandes matemáticos, como Emile Borel, no «creían» en las demostraciones de los teoremas de la forma: $(\exists x, A(x))$, que no proporcionan una construcción efectiva de una «solución» x que haga verdadera la relación $A(x)$.

En seguida se hizo patente la necesidad de introducir, en la teoría de conjuntos, un axioma llamado *axioma de la elección*; que grosso modo, asegura que siempre que una relación del tipo $\exists x, A(x)$ es verdadera, se puede *construir formalmente* un objeto x para el que $A(x)$ es verdadera.

Este axioma se puede formular de varias formas equivalentes, siendo la más conocida el *axioma de Zermelo* sobre conjuntos bien ordenados, y el *teorema de Zorn* (cf., por ejemplo, [7], o [4], Cap. I y II).

En la *matemática formal* usual, el axioma de la elección se introduce desde el principio con la ayuda de un signo lógico. En la teoría de conjuntos así construida, el símbolo $\exists x, A(x)$ no es más que una abreviatura para expresar de alguna forma que el objeto teórico que se puede construir y que verifica $A(x)$, la verifica efectivamente.

Ello no obsta para que se hayan construido otras teorías de conjuntos, en las que no se impone el axioma de la elección; por lo que es posible clasificar los resultados según que dependan o no del axioma de la elección.

Entre los resultados que dependen de este axioma, citemos, al azar: la existencia del producto de una familia cualquiera de conjuntos, la existencia de un ideal máximo en un anillo conmutativo, la existencia de una base en un espacio vectorial cualquiera, la existencia de una clausura algebraica para un cuerpo conmutativo cualquiera, la existencia de ultrafiltros. Estos objetos no pueden ser «construidos». Por ejemplo, no se conoce explícitamente un sistema S de números reales que constituya una base de \mathbf{R} , considerado como espacio vectorial sobre \mathbf{Q} . (A tal sistema se le llama una «base de Hamel».)

§ 1.4 OPERACIONES SOBRE LOS CONJUNTOS

Admitimos la noción de conjunto. Un conjunto E es, pues, un término, provisto de una relación: \in .

— La relación: $a \in E$ se lee « a pertenece a E » o « a es elemento de E »; la negación de esta relación se escribe: $a \notin E$.

Intuitivamente, E es la colección de los objetos a tales que $a \in E$, con exclusión de los otros.

Sea E un conjunto, y a, b dos elementos de E ; escribiremos $a = b$ para expresar que a y b son el mismo objeto (en realidad, esto no es absolutamente cierto; la noción de igualdad exigiría un análisis más profundo).

— Sean a, b, \dots, l, m objetos. Existe un conjunto E y sólo uno tal que

$$(x \in E) \Leftrightarrow ((x = a) \text{ o } (x = b) \dots \text{ o } (x = m)).$$

Este conjunto se designa $\{a, b, \dots, l, m\}$.

Inclusión

Sean E, F dos conjuntos. Escribiremos $E \subset F$ (E está contenido en F , o: E es una parte de F), para expresar la relación:

$$\forall x, (x \in E) \Rightarrow (x \in F).$$

Se tiene:

$$(E \subset F \text{ y } F \subset G) \Rightarrow (E \subset G),$$

y

$$(E \subset F \text{ y } F \subset E) \Leftrightarrow (E = F).$$

Construcción de conjuntos

No se dispone de ningún método efectivo para reconocer si un término dado es un conjunto; por lo que la teoría de conjuntos procede a su construcción, a partir de términos que se admiten como conjuntos (por ejemplo, \mathbf{N}). Vamos a pasar revista a los principales procedimientos para construir conjuntos.

1.º PARTE DE UN CONJUNTO DEFINIDA POR UNA RELACIÓN

Sean E un conjunto, y $A(x)$ una relación; a la relación $(x \in E \text{ y } A(x))$ se le llama *relación sobre E* . Admitiremos que existe un conjunto F tal que se verifica la equivalencia:

$$(\forall x, x \in F) \Leftrightarrow (x \in E \text{ y } A(x)) \quad (\text{la existencia de } F \text{ es un axioma}).$$

F está unívocamente determinado, y es evidentemente una parte de E , puesto que: $\forall x, (x \in F) \Rightarrow (x \in E)$. De ahora en adelante, en esta obra, designaremos a este conjunto F por

$$\{ x \mid x \in E \text{ y } A(x) \}.$$

Ejemplo

Sean E y F dos conjuntos tales que $E \subset F$. Al conjunto

$$\{ x \mid x \in F \text{ y } x \notin E \}$$

se le llama *complementario* de E en F , y se designa por $\mathbf{C}_F E$. Es claro que

$$\mathbf{C}_F(\mathbf{C}_F E) = E.$$

De ahora en adelante introduciremos la notación más cómoda $F \setminus E$ en vez de $\mathbf{C}_F E$.

Nota. Si $A(x)$ es una relación, en general, *no existe* ningún conjunto E que verifique la equivalencia:

$$(x \in E) \Leftrightarrow A(x).$$

Cuando tal conjunto existe, se dice que la relación $A(x)$ es *colectivizante en x* ; en caso contrario, es *no colectivizante*.

Por ejemplo, $x \notin E$ no es colectivizante en x , así como tampoco $x \notin x$.

TEOREMA I.4.1

|| Existe un conjunto, designado por \emptyset , tal que $\forall x, x \notin \emptyset$, y este conjunto es único.

A \emptyset se le llama *conjunto vacío*.

Demostración. Sea E un conjunto. Basta considerar $\emptyset = \mathbf{C}_E E$ (razonar por reducción al absurdo). Además, para cualquier otro conjunto F , se tiene $\emptyset \subset F$: ya que, al ser $(\forall x, x \notin \emptyset)$ verdadera, la relación $\forall x, (x \in \emptyset) \Rightarrow (x \in F)$ es verdadera. Ello prueba la unicidad de \emptyset . c.q.d.

2.º CONJUNTO DE LAS PARTES DE UN CONJUNTO

Sea E un conjunto. Se admite que existe un conjunto, designado $\mathcal{P}(E)$, tal que verifica la equivalencia

$$(X \in \mathcal{P}(E)) \Leftrightarrow (X \subset E) \text{ (la existencia de } \mathcal{P}(E) \text{ es un axioma).}$$

A $\mathcal{P}(E)$ se le llama el *conjunto de las partes de E*.

Dado que $\emptyset \in \mathcal{P}(E)$, vemos que $\mathcal{P}(E)$ es siempre no vacío.

3.º PRODUCTO CARTESIANO

Sean a y b dos términos; se admite que existe un término, designado (a, b) , que verifica la equivalencia:

$$((a', b') = (a, b)) \Leftrightarrow ((a = a') \text{ y } (b = b')). \quad (*)$$

A (a, b) se le llama el *par* de primer término a y segundo término b . La existencia del par es un axioma de la teoría de conjuntos.

Sean E y F dos conjuntos. Se admite que existe un conjunto, designado por $(E \times F)$, definido por la equivalencia:

$$(z \in E \times F) \Leftrightarrow (\exists a \in E, \exists b \in F, z = (a, b)).$$

(La existencia de $E \times F$ también constituye un axioma.)

A $(E \times F)$ se le llama **producto cartesiano** de E por F ; E es el *primer factor*, F es el *segundo factor* de $(E \times F)$. Por definición,

$$(E \times F = \emptyset) \Leftrightarrow (E = \emptyset \text{ o } F = \emptyset).$$

Cuando $F = E$, la **diagonal** de $E \times E$ es el conjunto de los pares (x, x) , con $x \in E$.

Se define asimismo la noción de producto cartesiano de un número finito n de conjuntos E, F, G, \dots, L, M , como el conjunto de las «n-plas» (a, b, c, \dots, l, m) , en donde $a \in E, b \in F, c \in G, \dots, l \in L, m \in M$. Este producto se designa por

$$E \times F \times G \times \dots \times L \times M.$$

4.º INTERSECCIÓN Y REUNIÓN DE DOS CONJUNTOS

Sean E y F dos conjuntos.

— Al conjunto:

$$\{ x \mid x \in E \text{ y } x \in F \}$$

(*) Por ejemplo, se puede tomar: $(a, b) = \{\{a, b\}, \{a\}\}$.

se le llama *intersección* de E y F , y se designa por $E \cap F$; es claro que su existencia es una consecuencia del procedimiento de construcción dado en 1°. La intersección verifica las siguientes propiedades:

$$\begin{aligned} E \cap F &= F \cap E && (\text{conmutatividad de } \cap), \\ (E \cap F) \cap G &= E \cap (F \cap G) && (\text{el resultado común se escribe } E \cap F \cap G: \\ &&& \text{asociatividad de } \cap), \\ E \cap \emptyset &= \emptyset, \\ (E \subset F) &\Leftrightarrow (E \cap F = E). \end{aligned}$$

— Se admite que la relación $((x \in E) \text{ o } (x \in F))$ es colectivizante en x (se obtiene de uno de los axiomas de la teoría de conjuntos). Al conjunto:

$$\{ x \mid (x \in E) \quad \text{o} \quad (x \in F) \}$$

se le llama *reunión* de E y F , y se designa por $E \cup F$. Se tienen las propiedades:

$$\begin{aligned} E \cup F &= F \cup E && (\text{conmutatividad}), \\ (E \cup F) \cup G &= E \cup (F \cup G) && (\text{el resultado común se escribe } E \cup F \cup G: \\ &&& \text{asociatividad de } \cup), \\ E \cup \emptyset &= E, \\ (E \subset F) &\Leftrightarrow (E \cup F = F). \end{aligned}$$

Finalmente, se tienen las siguientes propiedades, que relacionan la intersección y la reunión:

Si $E \subset G$ y $F \subset G$,

$$\begin{aligned} \mathbb{C}_G(E \cup F) &= (\mathbb{C}_G E) \cap (\mathbb{C}_G F), \\ \mathbb{C}_G(E \cap F) &= (\mathbb{C}_G E) \cup (\mathbb{C}_G F), \\ (E \cap F) \cup G &= (E \cup G) \cap (F \cup G) && (\text{distributividad de } \cup \text{ respecto de } \cap), \\ (E \cup F) \cap G &= (E \cap G) \cup (F \cap G) && (\text{distributividad de } \cap \text{ respecto de } \cup). \end{aligned}$$

5.° CONJUNTO COCIENTE: Ver más adelante.

6.° DIFERENCIA DE DOS CONJUNTOS

Sean E y F dos conjuntos. Según el apartado 1°, el conjunto

$$\mathbb{C}_E(E \cap F) = \{ x \mid (x \in E) \quad \text{y} \quad (x \notin F) \}$$

está bien definido. Este conjunto es la *diferencia de E y F* (o el complementario de F respecto de E). La notación: $(E \setminus F)$ tiende a extenderse.

Si $F \subset E$ se tiene que: $E \setminus F = \complement_E F$.

Relaciones entre reunión, intersección y producto

Sean E, F dos conjuntos. Si $A \subset E$ y $B \subset F$, se tiene: $A \times B \subset E \times F$, y recíprocamente. Supongamos que $A \subset E$ y $A' \subset E$. Es fácil comprobar las fórmulas:

$$(A \cup A') \times B = (A \times B) \cup (A' \times B),$$

$$(A \cap A') \times B = (A \times B) \cap (A' \times B),$$

y las fórmulas análogas para $B \subset F$, $B' \subset F$.

De ellas se deduce, por ejemplo

$$\begin{aligned} (A \cup A') \times (B \cap B') &= (A \times (B \cap B')) \cup (A' \times (B \cap B')) \\ &= [(A \times B) \cap (A \times B')] \cup [(A' \times B) \cap (A' \times B')]. \end{aligned}$$

Existencia de conjuntos

Los axiomas de la teoría de conjuntos implican, en particular, la existencia de *conjuntos finitos*. A partir de ellos es posible construir los *números enteros* y desarrollar, en parte, la teoría. (El lector puede consultar las obras de aritmética destinadas a las clases de terminal C, o [12].)

La existencia del *conjunto de los números enteros* (designado por \mathbf{N}) precisa de un nuevo axioma de la teoría de conjuntos, llamado el *axioma del infinito*, que se enuncia como sigue:

«Existe un conjunto infinito» (es decir, un conjunto que no es finito).

A partir del conjunto \mathbf{N} se construyen todos los conjuntos utilizados en las matemáticas usuales (cf. § 11).

§ I.5 CORRESPONDENCIAS Y APLICACIONES

DEFINICIÓN I.5.1

$\left\{ \begin{array}{l} \text{Sean } E \text{ y } F \text{ dos conjuntos. Se llama } \mathbf{grajo} \text{ de } E \text{ hacia } F \text{ a todo subcon-} \\ \text{junto de } E \times F. \text{ Se llama } \mathbf{correspondencia} \text{ de } E \text{ hacia } F \text{ a la terna} \\ (\Gamma, E, F), \text{ en donde } \Gamma \text{ es un grafo de } E \text{ hacia } F; E \text{ es el } \mathbf{conjunto de sa-} \\ \text{lida, } F \text{ es el } \mathbf{conjunto de llegada} \text{ de la correspondencia.} \end{array} \right.$

Sea $\gamma = (\Gamma, E, F)$ una correspondencia de E hacia F . Al conjunto (proyección de Γ sobre E)

$$\{ x \mid x \in E \text{ y } \exists y, (x, y) \in \Gamma \}.$$

se le llama *conjunto de definición* de γ .

El conjunto de definición de γ es vacío si, y sólo si, Γ es vacío.

Para toda parte A de E , se designa por $\gamma(A)$ el conjunto

$$\{ y \mid (y \in F) \text{ y } (\exists x, (x \in A \text{ y } (x, y) \in \Gamma)) \}.$$

Se tiene, pues, $\gamma(A) \subset F$.

Para toda parte B de F , se designa por $\gamma^{-1}(B)$ el conjunto

$$\{ x \mid (x \in E) \text{ y } (\exists y, (y \in B \text{ y } (x, y) \in \Gamma)) \}.$$

Se tiene, pues, $\gamma^{-1}(B) \subset E$.

Para simplificar la escritura para todo $x \in E$, se escribe:

$$\gamma(x) = \gamma(\{ x \})$$

y, para todo $y \in F$, se escribe:

$$\gamma^{-1}(y) = \gamma^{-1}(\{ y \}).$$

DEFINICIÓN I.5.2

$\left\{ \begin{array}{l} \text{Sea } E \text{ un conjunto. Se llama } \mathbf{relación\ binaria} \text{ en } E \text{ a toda correspon-} \\ \text{dencia de } E \text{ hacia } E. \end{array} \right.$

Sea R una relación binaria en E ; $R = (\Gamma, E, E)$. La relación: $x \in E, y \in E$ y $(x, y) \in \Gamma$ la designaremos, por convenio, $x R y$. Con este convenio, se hablará simplemente de «la relación $x R y$ en E ».

Ejemplos

1) La relación $x = x$ en E es una relación binaria, cuyo grafo es el conjunto Δ_E de los elementos de la forma (x, x) de $E \times E$; a este conjunto Δ_E se le llama la *diagonal* de $E \times E$.

2) Sea E un conjunto: la relación $A \subset B$ en $\mathcal{P}(E)$ es una relación binaria en $\mathcal{P}(E)$; si Γ designa la parte de $\mathcal{P}(E) \times \mathcal{P}(E)$ formada por los (A, B) tales que $A \subset B$,

la correspondencia $(\Gamma, \mathcal{P}(E), \mathcal{P}(E))$ admite a $\mathcal{P}(E)$ como dominio de definición, ya que $A \subset E$ para todo $A \in \mathcal{P}(E)$.

Más adelante estudiaremos dos tipos de relaciones particularmente importantes: las *relaciones de equivalencia* y las *relaciones de orden*. Observemos, finalmente, que los grafos de E hacia F constituyen un conjunto, que se identifica con $\mathcal{P}(E \times F)$.

No proseguiremos el estudio general de los grafos, y nos limitaremos a estudiar otra aplicación de esta noción, que es el *concepto de aplicación de un conjunto en otro conjunto*.

DEFINICIÓN I.5.3

Sean E y F conjuntos.

1) Se llama **grafo funcional** de E hacia F , a cualquier grafo Γ de E hacia F que verifique la relación siguiente:

Para todo $x \in E$, el conjunto $\{y \mid y \in F \text{ y } (x, y) \in \Gamma\}$
es vacío o está reducido a un elemento.

2) Se llama **aplicación**, o **función**, de E en F a toda correspondencia de E hacia F , cuyo grafo es un grafo funcional y cuyo dominio de definición es E .

La frase: « f es una aplicación de E en F », se escribe, por convenio:

$f: E \rightarrow F$. Se dirá también que « f toma sus valores en F ».

Al conjunto $f(E)$ se le llama *imagen* de f y también se designa por $\text{Im } f$.

Sea entonces $f: E \rightarrow F$, de grafo Γ . Con las notaciones que siguen a la definición I.5.1, se tiene que para todo elemento $x \in E$, $f(x)$ es un conjunto con un solo elemento, es decir, $f(x) = \{y\}$, para un cierto $y \in F$. Convenimos en designar al propio elemento y por medio de $f(x)$; y se le llama *transformado de x por f* ; cuando se quiere indicar en forma explícita el efecto de f sobre x , se escribe:

$f: E \rightarrow F, \quad x \mapsto f(x),$

y se lee: (f , aplicación de E en F , que envía x a $f(x)$).

Las aplicaciones de E en F constituyen un conjunto que se identifica con un subconjunto de $\mathcal{P}(E \times F)$ (ya que en § 4 anterior hemos identificado $\mathcal{P}(E \times F)$ con el conjunto de los grafos de E hacia F). A menudo a este conjunto lo designaremos

por $\mathcal{F}(E, F)$.

Evidentemente no se ha de confundir, $f(x)$ que es un elemento de F , con f que es un elemento de $\mathcal{F}(E, F)$.

Ejemplos de aplicaciones

- I_E , aplicación idéntica (*) de E en E , es tal que, para todo $x \in E$, $I_E(x) = x$.
- Si $F \subset E$, la inyección canónica j de F en E , es tal que, para todo $x \in F$, $j(x) = x$.

DEFINICIÓN I.5.4

Sean E, F, G conjuntos y $f: E \rightarrow F$, $g: F \rightarrow G$. Se llama **compuesta** de las aplicaciones g y f , y se designa por $g \circ f$, a la aplicación $h: E \rightarrow G$, tal que $h(x) = g(f(x))$ para todo $x \in E$.

Propiedad. Si $f: E \rightarrow F$, $g: F \rightarrow G$ y $h: G \rightarrow H$ son aplicaciones, se tiene:

$$(h \circ g) \circ f = h \circ (g \circ f) \text{ (asociatividad de la composición).}$$

El resultado común se designa por $h \circ g \circ f$.

Restricción y prolongación

Sean $f: E \rightarrow G$ y $g: F \rightarrow H$ dos aplicaciones, y $A \subset E \cap F$.

Diremos que f y g coinciden en A si se tiene:

$$f(x) = g(x) \text{ para todo } x \in A.$$

Hecha esta precisión, sea $f: E \rightarrow F$, y $A \subset E$. Llamaremos *restricción de f a A* a la aplicación de A en F que coincide con f en A . A esta restricción la designaremos por medio de f_A ; f_A está, pues, definida por la fórmula:

$$f_A(x) = f(x) \text{ para } x \in A; \quad f_A \text{ se designa frecuentemente por } f|_A.$$

Para toda pareja de conjuntos (X, Y) , designaremos por $\mathcal{F}(X, Y)$ el conjunto de las aplicaciones de X en Y . Volviendo a las anteriores notaciones es posible definir una aplicación, llamada *de restricción*:

$$\rho_A: \mathcal{F}(E, F) \rightarrow \mathcal{F}(A, F), \quad f \mapsto f_A.$$

(*) A esta aplicación también, se le llama, *aplicación identidad* de E . (N. del T.)

— Consideremos de nuevo $f: E \rightarrow G$ y $g: F \rightarrow G$. Diremos que g es una *prolongación de f* si se verifica que $E \subset F$, y que f y g coinciden en E . Por ejemplo, una aplicación es siempre una prolongación de una de sus restricciones.

«Dualmente» se define la noción de *correstricción* de una aplicación:

Si $f: E \rightarrow F$ es una aplicación, y si B es una parte de F tal que $\text{Im}(f) \subset B$, se llama *correstricción de f a B* a la aplicación

$$f^B: E \rightarrow B, \quad x \mapsto f^B(x) = f(x).$$

Es posible considerar el conjunto $\mathcal{F}^B(E, F)$ de las aplicaciones de E en F cuya imagen está contenida en B , y se tiene entonces una *biyección natural*, llamada de correstricción:

$$\mathcal{F}^B(E, F) \rightarrow \mathcal{F}(E, B), \quad f \mapsto f^B.$$

En la práctica, a fin de agilizar el lenguaje, el paso de una aplicación a una de sus correstricciones frecuentemente se sobreentiende.

DEFINICIÓN I.5.5

Sea $f: E \rightarrow F$ una aplicación.

1) f es **inyectiva** si se verifica la siguiente relación:

$$\forall x \in E, \forall y \in E, \quad x \neq y \Rightarrow f(x) \neq f(y).$$

2) f es **epiyectiva** si, para todo $y \in F$, existe un $x \in E$ tal que $f(x) = y$.

3) f es **biyectiva** si es, a la vez, inyectiva y epiyectiva.

A una biyección de un conjunto E en sí mismo se le llama, a veces, una *permutación de E* . (Esta terminología se utiliza sobre todo cuando E es finito.)

Una **involución** del conjunto E es una permutación j de E tal que $j \circ j = I_E$ y $j \neq I_E$.

I.5.1 La compuesta de dos aplicaciones inyectivas (resp. epiyectivas, biyectivas) es inyectiva (resp. epiyectiva, biyectiva) (evidente).

Aplicación recíproca

Designemos por f una aplicación biyectiva de E en F . Para todo $y \in F$, existe un elemento $x \in E$ y sólo uno tal que $f(x) = y$. Pongamos $x = f^{-1}(y)$. Se define así una aplicación $f^{-1}: F \rightarrow E$.

f^{-1} es también biyectiva, y para todo $x \in E$ se tiene: $f^{-1} \circ f(x) = x$
para todo $y \in F$ se tiene: $f \circ f^{-1}(y) = y$

Sea I_X la aplicación $x \mapsto x$ del conjunto X en sí mismo. A I_X se le llama la *aplicación idéntica de X* . Volviendo a la biyección $f: E \rightarrow F$, se tiene entonces:

$$(1) \quad f^{-1} \circ f = I_E \quad \text{y} \quad f \circ f^{-1} = I_F.$$

Además, $(f^{-1})^{-1} = f$. A f y f^{-1} se les llama *recíprocas* una de otra.

Vamos a establecer ahora una proposición recíproca de la I.5.1.

TEOREMA I.5.2

|| Sean $f: E \rightarrow F$ y $g: F \rightarrow G$ aplicaciones.
 || Si $g \circ f$ es inyectiva, f es inyectiva. Si $g \circ f$ es epiyectiva, g es epiyectiva.

Demostración. Supongamos que $g \circ f$ es inyectiva, y sean $x \in E$, $y \in E$ tales que $x \neq y$. Por hipótesis se tiene que $g \circ f(x) \neq g \circ f(y)$, es decir, $g(f(x)) \neq g(f(y))$, lo que implica $f(x) \neq f(y)$, luego f es inyectiva.

Supongamos que $g \circ f$ es epiyectiva, y sea $y \in G$. Por hipótesis, existe un $x \in E$ tal que $g \circ f(x) = y$, es decir, $g(f(x)) = y$; luego y es el transformado de $f(x)$ por g , lo que prueba que g es epiyectiva. c.q.d.

COROLARIO

|| Sea $f: E \rightarrow F$. Si existe $g: F \rightarrow E$ tal que $g \circ f = I_E$ y $f \circ g = I_F$, entonces f es biyectiva y $g = f^{-1}$.

Este corolario es extremadamente útil por cuanto permite demostrar que una aplicación es biyectiva con la ayuda de simples fórmulas acerca de compuestas de aplicaciones.

Imagen directa e imagen recíproca

Sea $f: E \rightarrow F$ una aplicación; volvamos a las notaciones que siguen a la definición I.5.1.

Para toda parte A de E , al conjunto $f(A)$ se le llama *imagen directa* de A por f .

Para toda parte B de F , al conjunto $f^{-1}(B)$ se le llama *imagen recíproca* de B por f .

(No se debe confundir la notación f^{-1} introducida aquí, con la notación f^{-1} introducida para indicar la recíproca de una biyección; en general, el contexto es lo suficientemente claro para que no haya temor de confusión.)

Se tiene pues:

$$f(A) = \{ y \mid y \in F \quad \text{y} \quad (\exists x, x \in A \quad \text{y} \quad f(x) = y) \} \quad \text{y}$$

$$f^{-1}(B) = \{ x \mid x \in E \quad \text{y} \quad f(x) \in B \}.$$

Se tienen las fórmulas siguientes (que no deben aprenderse de memoria, pero que deben ser reconocidas rápidamente en caso necesario):

- | | |
|---|--|
| (2) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$, | (6) $f(f^{-1}(B)) \subset B$, |
| (3) $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$, | (7) $f^{-1}(f(A)) \supset A$, |
| (4) $f(A \cup B) = f(A) \cup f(B)$, | (8) $f^{-1}(\mathbb{C}_F B) = \mathbb{C}_E(f^{-1}(B))$, |
| (5) $f(A \cap B) \subset f(A) \cap f(B)$. | |

A modo de ejemplo verificamos (2):

$f^{-1}(A \cap B)$ es el conjunto $\{x \mid x \in E \text{ y } f(x) \in A \cap B\}$,

es decir,

$$\{x \mid x \in E \text{ y } f(x) \in A \text{ y } f(x) \in B\}.$$

Por otra parte, $f^{-1}(A) \cap f^{-1}(B)$ es el conjunto

$$\{x \mid x \in E \text{ y } x \in f^{-1}(A) \text{ y } x \in f^{-1}(B)\},$$

es decir,

$$\{x \mid x \in E \text{ y } f(x) \in A \text{ y } f(x) \in B\}, \text{ de donde se sigue el resultado.}$$

Observemos que en (6) se puede substituir \subset por $=$ si f es epiyectiva, y que en (7) se puede substituir \supset por $=$ si f es inyectiva. Además:

I.5.3 Sea $f: E \rightarrow F$ una aplicación **inyectiva**. Si A y B son partes de E , se tiene:

$$\parallel f(A \cap B) = f(A) \cap f(B).$$

Demostración. Sea $E' = f(E)$. Designamos por g la aplicación de E en E' , que coincide con f en E ; g es una biyección y, para toda parte C de E , se tiene que $f(C) = g(C)$. Ello nos conduce, pues, a la fórmula (2), puesto que $g = (g^{-1})^{-1}$.]

Las nociones de imagen directa y de imagen recíproca permiten, dada $f: E \rightarrow F$, considerar aplicaciones «naturales» relacionadas con f , que ligan $\mathcal{P}(E)$ y $\mathcal{P}(F)$. De forma precisa, se definen

$$\begin{aligned} \hat{f}: \mathcal{P}(E) &\rightarrow \mathcal{P}(F) & \check{f}: \mathcal{P}(F) &\rightarrow \mathcal{P}(E), \\ A &\mapsto f(A) & B &\mapsto f^{-1}(B). \end{aligned}$$

Si se conviene en identificar, para cada $x \in E$, $\{x\}$ y x , entonces E se identifica con una parte de $\mathcal{P}(E)$; y entonces \hat{f} se presenta como una *prolongación* de f a $\mathcal{P}(E)$.

El lector verificará propiedades tales como « f es inyectiva $\Leftrightarrow \hat{f}$ es inyectiva», etcétera. (cf. ejercicios).

§ 1.6 FAMILIAS

Sean I y E dos conjuntos. A una aplicación de I en E , a veces, se le llama *familia* de elementos de E . (La única diferencia entre aplicación y familia es, pues, puramente psicológica...) Para designar una aplicación $i \mapsto f(i)$, cuando se habla de una familia, se prefiere utilizar la *notación de índices*

$$i \mapsto x_i.$$

Se habla entonces de la familia $(x_i)_{i \in I}$ de elementos de E . (Los «elementos» de la familia $(x_i)_{i \in I}$ son, pues, elementos de E y no la aplicación $i \mapsto x_i$. Pero el abuso de lenguaje que consiste en decir «la familia de elementos de E », no molesta.)

Sea $(x_i)_{i \in I}$ una familia de elementos de E , y sea J una parte de I . Por definición, la familia $(x_i)_{i \in J}$ es una *subfamilia* de $(x_i)_{i \in I}$ con índices en J .

La noción de familia permite extender considerablemente las operaciones de reunión, de intersección y de producto de conjuntos.

● Para simplificar, supondremos en todo momento que el conjunto I de índices es no vacío. Si I es finito, diremos que la familia es *finita*.

DEFINICIÓN 1.6.1

Sea E un conjunto, y $(A_i)_{i \in I}$ una familia de partes de E . Se llama *reunión* de la familia $(A_i)_{i \in I}$, y se designa por

$$\bigcup_{i \in I} A_i, \text{ al conjunto } \{ x \mid x \in E \text{ y } \exists i \in I, x \in A_i \}.$$

Se llama *intersección* de la familia $(A_i)_{i \in I}$, y se designa por

$$\bigcap_{i \in I} A_i, \text{ al conjunto } \{ x \mid x \in E \text{ y } \forall i \in I, x \in A_i \}.$$

Sea entonces \mathcal{G} una parte de $\mathcal{P}(E)$; \mathcal{G} se puede considerar como una familia de partes de E con la ayuda de la inyección canónica $j: A \mapsto A$ de \mathcal{G} en $\mathcal{P}(E)$. Entonces cada elemento de \mathcal{G} se tiene a sí mismo por índice, de forma que dicha familia se escribe $(j(A))_{A \in \mathcal{G}}$. (Nota: No confundir *parte de E* con *parte de $\mathcal{P}(E)$* .)

Esto establecido, se define la reunión y la intersección de los conjuntos de \mathcal{G} , como la reunión y la intersección de la familia definida anteriormente. Esta reunión y esta intersección se designan, respectivamente, por:

$$\bigcup_{A \in \mathcal{G}} A \quad \text{y} \quad \bigcap_{A \in \mathcal{G}} A.$$

Ejemplos

1) Si E es no vacío, se tiene

$$\bigcup_{x \in E} \{x\} = E, \quad \text{y} \quad \bigcap_{x \in E} \{x\} = \emptyset \text{ si } E \text{ tiene, por lo menos, dos elementos.}$$

2) Si E es no vacío, $\mathcal{P}(E) \setminus \{\emptyset\}$ es no vacío. Designando por $\mathcal{P}(E)^*$ a este conjunto, se tiene:

$$\bigcup_{A \in \mathcal{P}(E)^*} A = E, \quad \text{y si } E \text{ tiene por lo menos dos elementos} \quad \bigcap_{A \in \mathcal{P}(E)^*} A = \emptyset.$$

— Las fórmulas del § 4 (conmutatividad y asociatividad de \cup y \cap , distributividad de \cup respecto de \cap , y de \cap respecto de \cup) se generalizan de forma considerable.

TEOREMA I.6.1 (cambio de índices de una familia)

$$\left\| \begin{array}{l} \text{Sean } (A_i)_{i \in I} \text{ una familia de partes del conjunto } E; L, \text{ un conjunto no vacío,} \\ \text{y } f: L \rightarrow I \text{ una aplicación epiyectiva, en donde } I \text{ es no vacío. Entonces} \\ \text{se verifica:} \\ (1) \quad \bigcup_{\lambda \in L} A_{f(\lambda)} = \bigcup_{i \in I} A_i, \\ (2) \quad \bigcap_{\lambda \in L} A_{f(\lambda)} = \bigcap_{i \in I} A_i. \end{array} \right.$$

Demostración. Limitémonos a demostrar (2), y sean

$$V = \bigcap_{\lambda \in L} A_{f(\lambda)}, \quad W = \bigcap_{i \in I} A_i.$$

Para cualquier $i \in I$, existe un $\lambda \in L$ tal que $i = f(\lambda)$ (puesto que f es epiyectiva). Ahora bien, $x \in V$ implica $x \in A_{f(\lambda)}$; luego para todo $i \in I$, $x \in A_i$, de donde $V \subset W$.

Si $x \in W$, sean $\lambda \in L$ e $i = f(\lambda)$. Puesto que $x \in A_i$, se tiene que $x \in A_{f(\lambda)}$, para todo λ , de donde $W \subset V$, y $V = W$. c.q.d.

TEOREMA I.6.2 (Asociatividad)

Sea $(A_i)_{i \in I}$ una familia de partes del conjunto E . Se supone que $I = \bigcup_{\lambda \in L} J_\lambda$, con $L \neq \emptyset$ y $J_\lambda \neq \emptyset$ para todo $\lambda \in L$. Entonces se tiene:

$$(3) \quad \bigcup_{i \in I} A_i = \bigcup_{\lambda \in L} \left(\bigcup_{i \in J_\lambda} A_i \right),$$

$$(4) \quad \bigcap_{i \in I} A_i = \bigcap_{\lambda \in L} \left(\bigcap_{i \in J_\lambda} A_i \right).$$

Demostración inmediata

TEOREMA I.6.3 (Distributividad)

Sean $(A_i)_{i \in I}$ y $(B_j)_{j \in J}$ dos familias de partes de E . Si I y J son no vacíos, se tiene:

$$(5) \quad \left(\bigcap_{i \in I} A_i \right) \cup \left(\bigcap_{j \in J} B_j \right) = \bigcap_{(i,j) \in I \times J} (A_i \cup B_j),$$

$$(6) \quad \left(\bigcup_{i \in I} A_i \right) \cap \left(\bigcup_{j \in J} B_j \right) = \bigcup_{(i,j) \in I \times J} (A_i \cap B_j).$$

Demostración. Nos limitaremos, por ejemplo, a establecer (6). Hagamos

$$A = \bigcup_{i \in I} A_i, \quad B = \bigcup_{j \in J} B_j, \quad \text{y} \quad C = \bigcup_{(i,j) \in I \times J} (A_i \cap B_j).$$

Si $x \in A \cap B$, existe un i tal que $x \in A_i$, y un j tal que $x \in B_j$, de donde

$$x \in A_i \cap B_j \subset C.$$

Luego $A \cap B \subset C$; si $x \in C$, existe $(i, j) \in I \times J$ tal que $x \in A_i \cap B_j$, de donde $x \in A_i \subset A$ y $x \in B_j \subset B$, luego $x \in A \cap B$, lo que prueba que $C \subset A \cap B$, de donde, finalmente, $A \cap B = C$. c.q.d.

Estas fórmulas (5) y (6) admiten todavía una generalización.

A modo de ejercicio, el lector puede examinar en qué se transforman las fórmulas (2), (3), (4), (5) del § 5 cuando se substituyen $A \cap B$ y $A \cup B$ por intersecciones y uniones de familias cualesquiera de partes de E o de F .

Sucesiones

Sea E un conjunto. Por definición, una *sucesión de elementos de E* es una aplicación de \mathbf{N} en E . En el caso de las sucesiones se utiliza con preferencia la *notación de índices* en vez de la notación funcional, y se habla, por ejemplo, de la sucesión

$$n \mapsto u_n,$$

o, más simplemente, de la sucesión (u_n) ; u_n designa la imagen del entero n por la aplicación considerada.

No deberá confundirse una *sucesión* (u_n) con el *conjunto de valores* de dicha sucesión, que es la imagen de \mathbf{N} por esta aplicación. Por abuso de lenguaje, se permite, sin embargo, hablar del conjunto de los valores (u_n) , siempre que no induzca a confusión.

Una sucesión (u_n) de elementos de E se llama **estacionaria** si existe un entero n_0 tal que $u_n = u_{n_0}$ para $n \geq n_0$. El conjunto de valores de la sucesión es entonces finito; pero el recíproco es evidentemente falso: por ejemplo, el conjunto de los valores de la sucesión no estacionaria de enteros $(-1)^n$ se reduce a dos elementos.

Sea (u_n) una sucesión; por definición, una **subsucesión** de (u_n) es una sucesión de la forma $v_n = u_{\varphi(n)}$, en donde $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ es una aplicación estrictamente creciente.

§ I.7 PRODUCTOS

En lo que sigue, I es un conjunto no vacío.

Sea $(A_i)_{i \in I}$ una familia de partes del conjunto E ; si hacemos $A = \bigcup_{i \in I} A_i$, el conjunto A sólo depende de los A_i . Cuando no estemos interesados en E , se dirá simplemente:

«sea $(A_i)_{i \in I}$ una familia de conjuntos».

En estas condiciones, consideremos una familia $(x_i)_{i \in I}$ de elementos de A que está definida por la aplicación: $f : I \rightarrow A$, $i \mapsto x_i$. La f está determinada unívocamente por su grafo, que es un grafo funcional de I hacia A . En adelante, identificaremos la aplicación f con el grafo que la define; cuando se realiza esta identificación, f se convierte simplemente en una parte de $I \times A$.

DEFINICIÓN I.7.1

Sean $(A_i)_{i \in I}$ una familia de conjuntos y $A = \bigcup_{i \in I} A_i$. Se llama **producto** de la familia $(A_i)_{i \in I}$ al conjunto de las familias $(x_i)_{i \in I}$ de elementos de A tales que para todo $i \in I$, se verifica: $x_i \in A_i$. Se designa por $\prod_{i \in I} A_i$.

Si, para todo $i \in I$, $A_i \neq \emptyset$, entonces $\prod_{i \in I} A_i \neq \emptyset$. Esta propiedad constituye un axioma equivalente al *axioma de la elección*, del que ya hemos hablado. Nos dice que, si para todo i ($\exists a, a \in A_i$) es verdadera, se puede construir una aplicación $f: I \rightarrow A$ tal que, para cada i , $f(i) \in A_i$; lo que consiste en efectuar la «elección simultánea de todos los $f(i)$ uno en cada A_i ».

Si, para cada i , hacemos $A_i = E$ en la definición I.7.1 (siendo E un conjunto fijo), el producto $\prod_{i \in I} A_i$ se designa por E^I . Con los convenios que preceden a la definición I.7.1, vemos que es posible identificar E^I con el conjunto $\mathcal{F}(I, E)$ de las aplicaciones de I en E .

Volvamos al producto general $\prod_{i \in I} A_i = P$, y para cada i , consideremos la aplicación p_i :

$$p_i: P \rightarrow A_i,$$

en la que a cada $x = (x_j)_{j \in I}$ se le asocia el elemento x_i de A_i . A la aplicación p_i se le llama *la i -ésima proyección*.

La existencia de estas aplicaciones prueba que, si P es no vacío, ninguno de los A_i es vacío.

Finalmente, supongamos que el conjunto I es finito y está *ordenado* (cf. § 9); sea $I = \{i_1, i_2, \dots, i_n\}$. Entonces el producto $P = \prod_{i \in I} A_i$ es isomorfo al *producto cartesiano*

$$Q = A_{i_1} \times A_{i_2} \times \dots \times A_{i_n};$$

es decir, existe una biyección de P sobre Q . Esta biyección en la que al elemento $(x_i)_{i \in I}$ de P le corresponde el elemento $(x_{i_1}, x_{i_2}, \dots, x_{i_n})$ de Q se llama *canónica*. En este caso, al conjunto A_{i_k} se le puede llamar el *k -ésimo factor* de P (o de Q). Ordinariamente, P y Q no se distinguen.

Extensión de la noción de proyección

Sea $(A_{i \in I})$ una familia de conjuntos, y J una parte de I , no vacía; la aplicación:

$$p_J: \prod_{i \in I} A_i \rightarrow \prod_{i \in J} A_i,$$

$$(x_i)_{i \in I} \mapsto (x_i)_{i \in J},$$

que a toda familia $(x_i)_{i \in I}$, le hace corresponder su *restricción a J* , se denomina *proyección de índice J* . En virtud del axioma de la elección, p_J es epiyectiva si los A_i son no vacíos. En efecto, designemos por $X_J = (x_i)_{i \in J}$ un elemento de $\prod_{i \in J} A_i$; según

el axioma de la elección, existe una $f: \mathfrak{C}_I J \rightarrow \bigcup_{i \in I \setminus J} A_i$ tal que $f(i) \in A_i$, para todo $i \in I \setminus J$.

Hagamos $x_i = f(i)$ si $i \in I \setminus J$, $x_i = a_i$ si $i \in J$, y $x = (x_i)_{i \in I}$; se tiene entonces $p_J(x) = X_J$.

En particular, si los A_i son no vacíos, las $(p_i)_{i \in I}$ son epiyectivas, pues $p_i = p_{\{i\}}$ (si se identifican A_i y $\prod_{i \in \{i\}} A_i$).

No desarrollaremos las propiedades de los productos de conjuntos, y nos limitaremos al teorema siguiente:

TEOREMA I.7.1

Sean $(A_i)_{i \in I}$ una familia de conjuntos, y E un conjunto. Para toda aplicación $f: E \rightarrow \prod_{i \in I} A_i$, hagamos $p_i \circ f = f_i$. Por otra parte, para cada par de conjuntos (X, Y) designemos por $\mathcal{F}(X, Y)$ al conjunto de las aplicaciones de X en Y . Entonces, la aplicación $f \mapsto (f_i)_{i \in I}$ es una biyección de $\mathcal{F}(E, \prod_{i \in I} A_i)$ sobre $\prod_{i \in I} \mathcal{F}(E, A_i)$.

Demostración. Para toda $f \in \mathcal{F}(E, \prod_{i \in I} A_i)$ y todo $x \in E$, hacemos $f(x) = (y_i)_{i \in I}$.

Se tiene

$$y_i = p_i(f(x)) = p_i \circ f(x), \quad \text{pues} \quad f(x) = (p_i \circ f(x))_{i \in I},$$

luego f está unívocamente determinada por medio de las $p_i \circ f$. Por otra parte, si $(f_i)_{i \in I}$ es un elemento de $\prod_{i \in I} \mathcal{F}(E, A_i)$, la aplicación

$$f: E \rightarrow \prod_{i \in I} A_i, \quad x \mapsto (f_i(x))_{i \in I},$$

es tal que, para todo i , $f_i = p_i \circ f$. c.q.d.

§ 1.8 RELACIONES DE EQUIVALENCIA Y CONJUNTO COCIENTE

DEFINICIÓN 1.8.1

Sea R una relación binaria en un conjunto E ; R es una **relación de equivalencia** en E si verifica las siguientes propiedades:

E(1) para todo $x \in E$, se tiene: $x R x$ (reflexividad),

E(2) para todo $x \in E$ y todo $y \in E$, se tiene: $x R y \Rightarrow y R x$ (simetría),

E(3) para todo $x \in E$, todo $y \in E$ y todo $z \in E$, se tiene:

$(x R y \text{ e } y R z) \Rightarrow x R z$ (transitividad).

Consideremos, en el conjunto no vacío E , una relación de equivalencia R . Para todo $x \in E$, llamaremos *clase de equivalencia de x respecto de R* al conjunto $C(x) = \{y \mid y \in E \text{ y } x R y\}$; puesto que, según E_1 , $x \in C(x)$, $C(x) \neq \emptyset$. En virtud de E_2 , si $y \in C(x)$, se tiene: $x \in C(y)$. Por E_3 , si $y \in C(x)$, se tiene: $C(y) \subset C(x)$. De ello se deduce que, para todo $y \in C(x)$, se tiene: $C(y) = C(x)$.

Finalmente, sean x, y elementos cualesquiera de E ; en virtud de lo que antecede, se tiene $C(x) = C(y)$, o $C(x) \cap C(y) = \emptyset$.

DEFINICIÓN 1.8.2

Sea R una relación de equivalencia en el conjunto no vacío, E . Al subconjunto de $\mathcal{P}(E)$ formado por las clases de equivalencia respecto de R se le llama **conjunto cociente** de E por R , y se designa por: E/R .

Según las consideraciones precedentes a la definición 1.8.2, podemos enunciar:

TEOREMA 1.8.1

Sea R una relación de equivalencia en el conjunto no vacío, E . El conjunto cociente $\mathcal{S} = E/R$ posee las siguientes propiedades:

(P₁) $\forall x, (X \in \mathcal{S} \Rightarrow X \neq \emptyset)$,

(P₂) las relaciones $X \in \mathcal{S}$, $Y \in \mathcal{S}$ y $X \neq Y$ implican $X \cap Y = \emptyset$,

(P₃) $\bigcup_{X \in \mathcal{S}} X = E$.

DEFINICIÓN I.8.3

$\left\{ \begin{array}{l} \text{A toda parte } \mathcal{S} \text{ de } \mathcal{P}(E) \text{ que satisfaga } (P_1), (P_2), (P_3) \text{ se le llama parti-} \\ \text{ción de } E. \text{ En otras palabras, una } \textbf{partición} \text{ de } E \text{ es un conjunto de} \\ \text{partes de } E, \textbf{disjuntas y no vacías}, \text{ cuya reunión es } E. \end{array} \right.$

Por extensión, también se llama *partición* de E a toda familia $(X_i)_{i \in I}$ de partes no vacías de E que verifique:

- 1) $\forall i \in I, \forall j \in I, (i \neq j) \Rightarrow (X_i \cap X_j = \emptyset)$;
- 2) $\bigcup_{i \in I} X_i = E$.

El teorema I.8.1 admite un recíproco:

TEOREMA I.8.2

$\left\| \begin{array}{l} \text{Sea } \mathcal{S} \subset \mathcal{P}(E) \text{ una partición del conjunto } E; \text{ existe una relación de equi-} \\ \text{valencia } R \text{ en } E, \text{ única, tal que } \mathcal{S} = E/R. \end{array} \right.$

Demostración. Si R existe, las clases respecto de R son necesariamente los elementos de \mathcal{S} . Luego R debe ser la relación:

$$(1) \quad \exists X, X \in \mathcal{S}, x \in X \text{ y } y \in X.$$

Llamemos R a la relación binaria (1) (entre x e y). En virtud de (P_3) , R verifica E_1 ; E_2 es evidente; y E_3 resulta de (P_2) . En cuanto a (P_1) , sirve para comprobar que se tiene exactamente $E/R = \mathcal{S}$ c.q.d.

Aplicación canónica

Sea R una relación de equivalencia en el conjunto E ; según lo que precede, para todo $x \in E$, existe un único elemento X de E/R tal que $x \in X$; X es, precisamente, la clase de equivalencia de x respecto de R . Hagamos $X = p(x)$.

A la aplicación

$$p : E \rightarrow E/R, \quad x \mapsto p(x)$$

se le llama la *aplicación* (o: la *proyección*) *canónica* de E en E/R .

Observemos que la relación: $x R y$ equivale a: $p(x) = p(y)$, y que p es epiyectiva. Se tiene entonces la siguiente propiedad «universal»:

TEOREMA I.8.3

Sea R una relación de equivalencia en E , $p : E \rightarrow E/R$ la aplicación canónica, y $f : E \rightarrow F$ una aplicación. Si f es constante sobre las clases de equivalencia respecto de R , existe una aplicación $\bar{f} : E/R \rightarrow F$, única, tal que $\bar{f} \circ p = f$.

Demostración. Sea $X \in E/R$. Por hipótesis, si $x \in X$ e $y \in X$, $f(x) = f(y)$. Entonces hacemos $\bar{f}(X) =$ valor común de los $f(x)$ para $x \in X$. El resto del teorema es evidente. c.q.d.

TEOREMA I.8.4 (Descomposición canónica de una aplicación)

Sea $f : E \rightarrow F$ una aplicación. Designemos por R la relación binaria en E definida por:

$$x R y \Leftrightarrow f(x) = f(y).$$

Entonces R es una relación de equivalencia; si $j : f(E) \rightarrow F$ designa la inyección canónica, y si $p : E \rightarrow E/R$ es la aplicación canónica, existe una única aplicación $\bar{f} : E/R \rightarrow f(E)$ tal que: $j \circ \bar{f} \circ p = f$. Además, \bar{f} es una biyección.

$$E \xrightarrow{p} E/R \xrightarrow{\bar{f}} f(E) \xrightarrow{j} F$$

Demostración. Se verifica que R es una relación de equivalencia. La existencia y la unicidad de \bar{f} están aseguradas por el teorema I.8.3. Puesto que $\bar{f} \circ p$ coincide con f en E , $\bar{f} \circ p$ es epiyectiva, luego \bar{f} es epiyectiva. Demostremos, finalmente, que \bar{f} es inyectiva: elijamos X e Y en E/R tales que $\bar{f}(X) = \bar{f}(Y)$, y sean $x \in X$ e $y \in Y$. Se tiene: $f(x) = \bar{f}(X) = \bar{f}(Y) = f(y)$, luego $x R y$, luego $X = Y$. c.q.d.

A la descomposición $f = j \circ \bar{f} \circ p$ se le llama *descomposición canónica* o *factorización canónica*, de f . La sucesión de aplicaciones dada anteriormente queda mejor «visualizada» por medio del diagrama cuadrado que sigue:

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ p \downarrow & & \uparrow j \\ E/R & \xrightarrow{\bar{f}} & f(E) \end{array}$$

La relación $f = j \circ \bar{f} \circ p$ se expresa diciendo que este diagrama es *conmutativo*. (i.e., se pueden seguir las flechas que se quiera para ir de E a F .)

Relación de equivalencia inducida

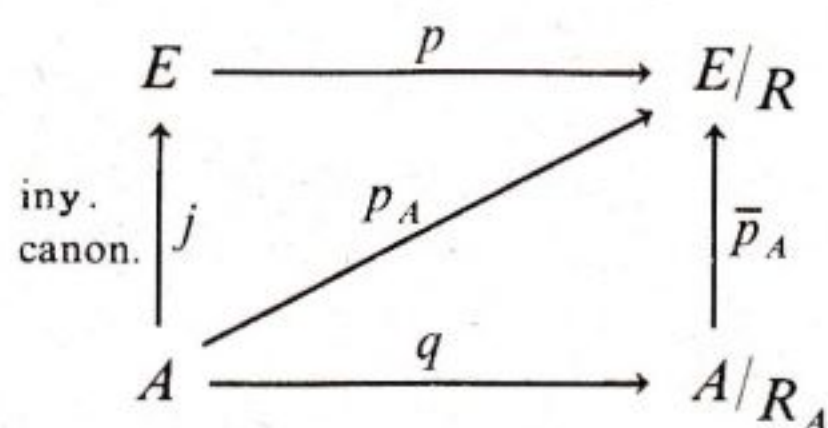
Sea R una relación de equivalencia en E , y A una parte no vacía de E . Designemos por R_A la relación:

$$x \in A \quad y \quad y \in A \quad y \quad x R y.$$

Entonces R_A es una relación de equivalencia en A , llamada *relación inducida en A por R* . Sea \mathcal{S}_A la parte de E/R formada por los $X \in E/R$ tales que $X \cap A \neq \emptyset$. Entonces A/R_A es el conjunto formado por los $X \cap A$, con $X \in \mathcal{S}_A$.

Sea $p: E \rightarrow E/R$ la aplicación canónica. La restricción p_A de p en A es constante sobre las clases de equivalencia respecto de R_A , luego p_A se factoriza:

$p_A = \bar{p}_A \circ q$, de donde $q: A \rightarrow A/R_A$, es la aplicación canónica. Se tiene, pues, el siguiente diagrama conmutativo:



en el cual $\bar{p}_A \circ q = p_A = p \circ j$.

Luego \bar{p}_A es una *inyección*, pues toda clase respecto de R_A está contenida en una clase respecto de R . A la aplicación \bar{p}_A se le llama *inyección canónica* (de A/R_A en E/R).

Ejemplos

1) Sean E y F dos conjuntos, A una parte de E , $\mathcal{F}(E, F)$ el conjunto de las aplicaciones de E en F . Sea R la relación binaria sobre $\mathcal{F}(E, F)$ definida por

$$(fRg) \Leftrightarrow (f(x) = g(x) \text{ para todo } x \in A)$$

R es una relación de equivalencia. Consideremos la aplicación

$$\rho: \mathcal{F}(E, F) \rightarrow \mathcal{F}(A, F),$$

que transforma a $f \in \mathcal{F}(E, F)$ en su *restricción* a A . ρ es constante sobre las clases respecto de R , luego se factoriza en la forma $\rho = \bar{\rho} \circ p$, en donde

$$p : \mathcal{F}(E, F) \rightarrow \mathcal{F}(E, F)/R$$

es la aplicación canónica. Se ve fácilmente que

$$\bar{\rho} : (\mathcal{F}(E, F))/R \rightarrow \mathcal{F}(A, F) \text{ es una biyección.}$$

2) Sean E un conjunto y A una parte de E , y R_A la relación binaria definida en $\mathcal{P}(E)$ por:

$$(X R_A Y) \Leftrightarrow (X \cup A = Y \cup A);$$

R_A es una relación de equivalencia. La aplicación $\rho_A : \mathcal{P}(E) \rightarrow \mathcal{P}(E \setminus A)$ tal que $\rho_A(X) = X \cap (E \setminus A)$ es constante sobre las clases respecto de R_A . Si

$$\mathcal{P}(E) \xrightarrow{p} \mathcal{P}(E)/R_A \xrightarrow{\bar{\rho}_A} \mathcal{P}(E \setminus A)$$

es la factorización canónica de ρ_A , $\bar{\rho}_A$ es una biyección.

Si F es una parte cualquiera de E , la relación inducida por R_A en $\mathcal{P}(F)$ es (para el conjunto F) la relación $R_{A \cap F}$.

— En el estudio de las estructuras algebraicas encontraremos gran número de relaciones de equivalencia. Estas relaciones sirven también para definir, a partir de conjuntos dados, nuevos conjuntos.

§ 1.9 RELACIONES DE ORDEN

DEFINICIÓN 1.9.1

1) Sea R una relación binaria en un conjunto E . Diremos que R es una **relación de orden en E** si se verifican las siguientes propiedades:

(O₁) para todo $x \in E$, $x R x$ (reflexividad);

(O₂) para todo $x, y, z \in E$, $((x R y) \text{ e } (y R z)) \Rightarrow (x R z)$ (transitividad);

(O₃) para todo $x, y \in E$, se tiene

$((x R y) \text{ e } (y R x)) \Rightarrow (x = y)$ (antisimetría).

2) Un conjunto provisto de una relación de orden se llama **conjunto ordenado**.

Corrientemente, una relación de orden se designa por \leq . En este caso, la relación « $x \leq y$ e $y \neq x$ » se designa por: « $x < y$ ». Cuando nos refiramos a una relación de orden sin precisar la notación, entenderemos que se designa por \leq .

Un orden es *total* si $(x \neq y) \Rightarrow ((x < y) \text{ o } (y < x))$; *parcial* en caso contrario.

Dos elementos x, y de un conjunto parcialmente ordenado son *comparables* si se verifica $x \leq y$ o $y \leq x$.

Ejemplos

1) Sea E un conjunto, en el conjunto $\mathcal{P}(E)$, la relación $X \subset Y$ es una relación de orden, y en general, en toda parte \mathcal{S} de $\mathcal{P}(E)$, la relación $X \subset Y$ es una relación de orden. Cuando en \mathcal{S} se considera esta relación, se dice que \mathcal{S} está *ordenado por inclusión*. Si E tiene por lo menos dos elementos, $\mathcal{P}(E)$ *no está* totalmente ordenado por inclusión.

2) Sea E un conjunto ordenado, y sea A un conjunto cualquiera. En el conjunto $\mathcal{F}(A, E)$ de las aplicaciones de A en E , la relación $f \leq g$, definida por:

$$(f \leq g) \Leftrightarrow (\forall x, (x \in A)) \Rightarrow (f(x) \leq g(x))$$

es una relación de orden; sólo cuando A se reduce a un elemento resulta que este orden es total.

3) Sea \mathbf{N}^* el conjunto de los enteros ≥ 1 ; designemos por $x|y$ la relación: « x divide a y »; $x|y$ es una relación de orden parcial en \mathbf{N}^* .

4) En el conjunto \mathbf{N} de los enteros naturales, la relación habitual de desigualdad $x \leq y$, que se puede definir por

$$\exists z, \quad z \in \mathbf{N} \quad \text{y} \quad y = x + z,$$

es una relación de orden. Se llama *orden natural* de \mathbf{N} . Es total.

5) Sean E un conjunto y Δ el conjunto de las particiones de E ; (Δ también es un conjunto, y es una parte de $\mathcal{P}(\mathcal{P}(E))$, o, si se prefiere, un elemento de $\mathcal{P}(\mathcal{P}(\mathcal{P}(E)))$). Dadas dos particiones \mathcal{S} y \mathcal{T} de E se dice que « \mathcal{S} es más fina que \mathcal{T} » si, para todo $S \in \mathcal{S}$, existe un $T \in \mathcal{T}$ tal que $S \subset T$. Equivale a decir que cada $T \in \mathcal{T}$ es reunión de elementos de \mathcal{S} .

Entonces la relación « \mathcal{S} es más fina que \mathcal{T} » es una relación de orden en Δ .

Orden inducido

Sea E un conjunto ordenado por \leq y sea A una parte de E . La relación

$$x \in A, \quad y \in A \quad \text{y} \quad x \leq y$$

es una relación de orden en A , llamada *relación de orden inducida* por E en A . Si A está provisto del orden inducido, se le llama *subconjunto ordenado* de E .

Aplicaciones monótonas

Si E y F designan conjuntos ordenados, una aplicación $f: E \rightarrow F$ es *creciente* si las relaciones $x \in E$, $y \in E$ y $x \leq y$ implican $f(x) \leq f(y)$; *decreciente* si dichas relaciones implican $f(x) \geq f(y)$.

$f: E \rightarrow F$ es *estrictamente creciente* si las relaciones $x \in E$, $y \in E$, y $x < y$ implican $f(x) < f(y)$; *estrictamente decrecientes* si dichas relaciones implican $f(x) > f(y)$.

Ejemplos

1) Ordenemos $\mathcal{P}(E)$ por inclusión; la aplicación $A \mapsto \mathbf{C}_E A$ es estrictamente decreciente.

2) Sea $f: E \rightarrow F$ una aplicación. Ordenemos $\mathcal{P}(E)$ y $\mathcal{P}(F)$ por inclusión; la aplicación

$$\dot{f}: \mathcal{P}(F) \rightarrow \mathcal{P}(E)$$

$$B \mapsto f^{-1}(B) \text{ es creciente.}$$

3) Sea E el conjunto de los enteros ≥ 1 , ordenado por « x divide a y ». Para todo entero $\alpha \geq 1$, la aplicación $x \mapsto x^\alpha$ es creciente.

Ultimo elemento. Primer elemento

DEFINICIÓN I.9.2

$\left\{ \begin{array}{l} \text{Sea } E \text{ un conjunto ordenado. Un elemento } m \in E \text{ se llama } \mathbf{primer} \\ \mathbf{elemento} \text{ de } E \text{ si, para todo } x \text{ de } E, \text{ se tiene } x \geq m; \text{ y a } M \in E \text{ se le} \\ \text{llama } \mathbf{último elemento} \text{ de } E \text{ si, para todo } x \in E, \text{ se tiene } x \leq M. \end{array} \right.$

Un primer elemento también se llama *elemento mínimo*.

Un último elemento también se llama *elemento máximo*.

En virtud de (O_3) , si E admite un primer (resp. último) elemento, éste es *único*. Se puede, pues, decir que es *el* primer (resp. último) elemento.

Ejemplos

1) Si $\mathcal{P}(E)$ está ordenado por inclusión (E , conjunto), \emptyset es el primer elemento, y E es el último elemento de $\mathcal{P}(E)$.

2) El conjunto \mathbf{N} de los enteros naturales, ordenado por el orden natural, admite a 0 como primer elemento, pero carece de último elemento.

Toda parte no vacía de \mathbf{N} (ordenada por el orden natural) admite primer elemento: esta propiedad de \mathbf{N} es fundamental.

3) El conjunto E de los enteros > 1 , ordenado por « x divide a y » carece de primero y último elementos.

Supremo e ínfimo

En lo que sigue, vamos a definir y estudiar la noción de *cota superior* (o *mayorante*) y de *supremo*. Las definiciones y propiedades de las cotas inferiores (o *minorantes*) y de los ínfimos se obtienen invirtiendo los signos de las desigualdades, y no las formularemos. (¡Pero las utilizaremos!).

DEFINICIÓN I.9.3

Sean E un conjunto ordenado, y A una parte de E . Un elemento $M \in E$ es una **cota superior** de A si, para todo $a \in A$, se tiene:

$$a \leq M.$$

Si existe, por lo menos, una cota superior de A , se dice que A está **acotado superiormente**. (*)

Es conveniente fijarse en el hecho de que la propiedad « A está acotado superiormente» es una propiedad del par (A, E) . En rigor se debería decir « A está acotado superiormente en E ». Sin embargo, a no ser que tratemos varios conjuntos ordenados y que A sea un subconjunto ordenado de todos ellos, la anterior precisión se omitirá.

Ejemplos

1) Sea E un conjunto, y sea $\mathcal{P}(E)$ ordenado por inclusión; el conjunto de partes de E reducidas a un elemento está acotado superiormente en $\mathcal{P}(E)$, pero, en

(*) El original francés utiliza los términos de mayorante (resp. minorante) y de conjunto mayorado (resp. minorado). Si bien podríamos haberlos adoptado en la traducción, hemos preferido utilizar los de cota superior (resp. cota inferior) y conjunto acotado superiormente (resp. acotado inferiormente) por cuanto son más corrientes en esta clase de literatura. (N. del T.)

cambio, si E tiene, por lo menos, dos elementos, no está acotado superiormente en $\mathcal{P}(E) \setminus \{E\}$.

2) Toda parte de \mathbf{N} no vacía y acotada superiormente, admite último elemento.

DEFINICIÓN I.9.4

*Sea E un conjunto ordenado, y sea A una parte de E . El **supremo de A en E** es el primer elemento (si existe) del conjunto de las cotas superiores de A en E ; se le designa entonces por $\sup_{x \in A} x$, o $\sup x$ si no hay peligro de confusión. (*)*

También aquí, el supremo, cuando existe, es una propiedad del par (A, E) .

Ejemplos

1) Sea E el conjunto de los enteros ≥ 1 , ordenado por $x|y$ (x divide a y). Toda parte finita A , no vacía, de E posee un supremo, que es el mcm de los elementos de A . Una parte no vacía, A , de E posee un ínfimo, que es el mcd de los elementos de A .

2) Sea E un conjunto y $\mathcal{P}(E)$ ordenado por inclusión. Sea \mathcal{S} una parte no vacía de $\mathcal{P}(E)$; entonces \mathcal{S} posee en $\mathcal{P}(E)$ un supremo, que es $\bigcup_{X \in \mathcal{S}} X$, y un ínfimo, que es $\bigcap_{X \in \mathcal{S}} X$.

3) Sea \mathbf{Q} el conjunto de los números racionales y \mathbf{R} el conjunto de los números reales; ordenamos \mathbf{R} mediante la relación de orden natural; \mathbf{Q} es un subconjunto ordenado de \mathbf{R} ; sea, entonces, el conjunto A

$$\{x \mid x \in \mathbf{Q}, x \geq 0 \text{ y } x^2 \leq 2\}.$$

A posee un supremo en \mathbf{R} , que es $\sqrt{2}$; pero A carece de supremo en \mathbf{Q} ; sea E el subconjunto de \mathbf{R} dado por $\mathbf{Z} \cup A$; entonces A posee un supremo en E , que es 2.

Caracterización del supremo

Según la definición I.9.5, para que un elemento b sea el supremo de A en E , es necesario y suficiente que se verifiquen las dos condiciones siguientes;

1) b es una cota superior de A , es decir, para todo $a \in A$, se tiene: $a \leq b$.

(*) El original francés utiliza el término «la cota superior de A en E » para referirse al supremo. (Nota del Traductor.)

2) b es la menor de las cotas superiores de A en E ; esto significa: si $c \in E$ es tal que no verifica $b \leq c$, entonces existe un elemento $a \in A$ que no verifica $a \leq c$.

Cuando E es totalmente ordenado, la condición 2) se puede escribir:

2 bis) Para todo elemento $c \in E$ tal que $c < b$, existe $a \in A$ tal que $a > c$.

Dada la importancia de este resultado, enunciaremos:

TEOREMA I.9.1

Sea A una parte de un conjunto **totalmente ordenado** E . Para que un elemento $b \in E$ sea el supremo de A en E , es necesario y suficiente que se verifiquen las dos condiciones siguientes:

- 1) para todo $a \in A$, se tiene: $a \leq b$.
- 2) para todo elemento $c \in E$ tal que $c < b$, existe $a \in A$ tal que $a > c$.

Supremo de una función con valores en un conjunto ordenado

Consideremos un conjunto ordenado E , un conjunto A , y una aplicación $f: A \rightarrow E$.

Si $\sup_{y \in f(A)}$ existe, a este elemento se le llama *supremo de f en E* , y se designa por $\sup_{x \in A} f(x)$.

Elemento maximal (resp. minimal)

DEFINICIÓN I.9.5

Sea E un conjunto ordenado y m un elemento de E . Se dice que m es **maximal** si las relaciones $x \in E$ y $x \geq m$ implican $x = m$.

— En otras palabras, m es maximal si $\{m\}$ no admite cota superior estricta. Se define análogamente un elemento minimal de E : m es minimal si $\{m\}$ no admite una cota inferior estricta, es decir, si: $x \in E$ y $x \leq m$ implican $x = m$.

— Si E posee elemento máximo M , M es maximal, y es el único elemento maximal de E . Pero si E carece de elemento máximo, puede perfectamente poseer *elementos maximales*. Sobre todo en este último caso es cuando la noción de elemento maximal adquiere toda su importancia.

— Observemos que si E es *totalmente* ordenado, las nociones de elemento máximo y maximal coinciden.

Ejemplos

1) Sea E^* el conjunto de los enteros > 1 , ordenado por « x divide a y ». E^* carece de mínimo (puesto que el único número que divide a todos los enteros es 1). Sin embargo, E^* posee una infinidad de elementos minimales, que son los números primos.

2) (Cf. Cap. VIII.) Sea E un espacio vectorial de dimensión n sobre un cuerpo K . Designamos por $\mathcal{G}(E)$ al conjunto de los subespacios de E distintos de E , y ordenamos $\mathcal{G}(E)$ por inclusión: $\mathcal{G}(E)$ carece de máximo, pero posee elementos maximales, que son los subespacios de dimensión $n - 1$ (hiperplanos vectoriales).

§ I.10 ENUMERACION

Suponemos conocidas todas las definiciones y propiedades elementales relativas a los números enteros naturales. A lo largo de toda esta obra, al conjunto de los números enteros se le designará por \mathbf{N} .

Recordemos la propiedad fundamental de \mathbf{N} :

TEOREMA I.10.1

|| Toda parte no vacía de \mathbf{N} admite un primer elemento (para el orden natural), y toda parte no vacía de \mathbf{N} , acotada superiormente, admite un último elemento.

Esta propiedad implica el teorema de recurrencia, que es uno de los medios más importante del razonamiento matemático.

TEOREMA I.10.2

|| Sea $P(n)$ una relación que dependa del entero n ; si $P(0)$ es verdadera, y si, para todo n , la relación $P(n) \Rightarrow P(n + 1)$ es verdadera, entonces $P(n)$ es verdadera para todo entero n .

Demostración. Razonemos por reducción al absurdo y supongamos que el conjunto A de los enteros n para los que $P(n)$ no es verdadero, es no vacío. Según

el teorema I.10.1, A posee un primer elemento m ; puesto que $P(0)$ es verdadera, se tiene que: $m \geq 1$; por definición de m , se tiene: $m - 1 \notin A$ ⁽¹⁾; luego $P(m - 1)$ es verdadera. Puesto que $P(m - 1) \Rightarrow P(m)$ es verdadera, se deduce que $P(m)$ es verdadera, lo que es absurdo. c.q.d.

Conjuntos finitos

Sea \mathbf{N}_n^* el conjunto $\{1, 2, \dots, n\}$ de los n primeros enteros no nulos.

Un conjunto A es *finito* si existe un entero n para el cual es posible encontrar una biyección de \mathbf{N}_n^* sobre A ; n es entonces único, y este entero se llama el *número de elementos de A* , o *cardinal de A* ; a menudo se le designa $\text{card}(A)$.

Si $f: E \rightarrow F$ es una aplicación de un conjunto finito en un conjunto finito, y si $\text{card}(E) = \text{card}(F)$, se tienen las equivalencias:

$$(f \text{ biyectiva}) \Leftrightarrow (f \text{ inyectiva}) \Leftrightarrow (f \text{ epiyectiva}).$$

Se demuestra que cada una de estas propiedades *caracteriza* a los conjuntos finitos.

Si A es un conjunto no finito, se dice que es *infinito*: por ejemplo, \mathbf{N} es infinito. Se demuestra, entonces, que existe una aplicación *inyectiva* de \mathbf{N} en A .

Recordemos, brevemente, algunas propiedades de los conjuntos finitos.

TEOREMA I.10.3

$$\left\| \begin{array}{l} \text{Sean } E \text{ y } F \text{ conjuntos finitos, y sea } \mathcal{F}(E, F) \text{ el conjunto de las aplicaciones} \\ \text{de } E \text{ en } F. \text{ Entonces } \mathcal{F}(E, F) \text{ es finito, y} \\ \text{card}(\mathcal{F}(E, F)) = p^n \text{ donde } n = \text{card}(E) \text{ y } p = \text{card}(F). \end{array} \right.$$

Demostración. Por recurrencia sobre n . Para $n = 1$ es evidente. Supongamos la propiedad verdadera para $\text{card}(E) = n - 1$, y demostrémosla para $\text{card}(E) = n$; entonces se tiene:

$$E = E' \cup \{a\}, \text{ donde } \text{card}(E') = n - 1.$$

Una aplicación de E en F está unívocamente determinada por su restricción a E' y por la imagen de a . Existen p imágenes posibles para a , y p^{n-1} restricciones posibles a E' por hipótesis de recurrencia. El número de aplicaciones de E en F es, pues, $p^{n-1} \times p = p^n$. c.q.d.

⁽¹⁾ La existencia de $m-1$ puede considerarse una consecuencia de la segunda afirmación del teorema I.10.1.

COROLARIO

|| Si E es un conjunto finito, y si $\text{card}(E) = n$, $\mathcal{P}(E)$ es finito y $\text{card}((\mathcal{P}(E))) = 2^n$.

Demostración. Sea $F = \{0,1\}$. A cada parte A de E hagámosle corresponder la aplicación $\chi_A : E \rightarrow F$ (llamada función característica de A) tal que $\chi_A(x) = 1$ si $x \in A$ y $\chi_A(x) = 0$ si $x \in E \setminus A$. Entonces la aplicación $A \mapsto \chi_A$ es una biyección de $\mathcal{P}(E)$ sobre $\mathcal{F}(E, F)$, y con la ayuda del teorema I.10.3, el teorema queda demostrado. ||

TEOREMA I.10.4

|| Sean E y F conjuntos finitos, y sean $m = \text{card}(E)$, $n = \text{card}(F)$, $n \geq m$. El número de aplicaciones **inyectivas** de E en F es entonces

$$n(n-1)(n-2)\dots(n-m+1) = \frac{n!}{(n-m)!}.$$

A este número se le llama, a veces, «número de variaciones de n objetos, tomados de m en m », y se designa frecuentemente por A_n^m .

Demostración. Por recurrencia sobre m , siendo el teorema evidente para $m = 1$.

Supongamos el teorema verdadero para $\text{card } E = m-1$, y probemos que si $m \leq n$, es entonces verdadero para $\text{card } E = m$; se tiene, pues $E = E' \cup \{a\}$, con $\text{card } E' = m-1$. Una aplicación $f : E \rightarrow F$ está unívocamente determinada por la elección de $f(a)$ y por la restricción f' de f a E' . Si f es inyectiva, f' es inyectiva; si f' es inyectiva, f es inyectiva si, y sólo si, $f(a) \notin f(E')$. Luego, fijado f' , existen $n - m + 1$ posibilidades para $f(a)$ ya que $\text{card } f'(E') = m-1$. Puesto que, por hipótesis de recurrencia, existen

$$\frac{n!}{(n-m+1)!}$$

posibilidades para f' , existen $\frac{n!}{(n-m+1)!} \times (n-m+1) = \frac{n!}{(n-m)!}$ posibilidades para f . c.q.d.

COROLARIO

|| El número de biyecciones de un conjunto de n elementos sobre sí mismo (**permutaciones** de este conjunto) es $n!$

TEOREMA I.10.5

Sea E un conjunto finito, y pongamos $n = \text{card}(E)$, y sea p un entero $\leq n$.
El número de partes A de E tales que $\text{card}(A) = p$ es igual a:

$$\frac{n!}{p!(n-p)!} = \frac{n(n-1)(n-2)\dots(n-p+1)}{p!}.$$

Demostración. Designemos por \mathbf{N}_p^* al conjunto $\{1, 2, \dots, p\}$ de los p primeros números enteros no nulos, y por \mathcal{O}_p al conjunto de las aplicaciones inyectivas de \mathbf{N}_p^* en E . Sabemos que $\text{card}(\mathcal{O}_p) = \frac{n!}{(n-p)!}$.

En \mathcal{O}_p , la relación R definida por $(f R g) \Leftrightarrow (f(\mathbf{N}_p^*) = g(\mathbf{N}_p^*))$ es una relación de equivalencia. Para todo elemento F del conjunto cociente \mathcal{O}_p/R , designamos por $\varphi(F)$ a la imagen común de los $f \in F$.

Por definición, la aplicación $\varphi : \mathcal{O}_p/R \rightarrow \mathcal{P}(E)$ es una inyección. Por otra parte la imagen de φ es el conjunto $\mathcal{P}_p(E)$ de las partes de E de p elementos; puesto que, para todo $A \in \mathcal{P}_p(E)$, existe una biyección f de \mathbf{N}_p^* en A . Luego φ es una biyección de \mathcal{O}_p/R en $\mathcal{P}_p(E)$.

Fijemos una parte A de E con p elementos: es claro que el número de elementos de $\varphi^{-1}(A)$, es igual al número de biyecciones de \mathbf{N}_p^* en A . Según el teorema I.10.4, este número es $p!$. Hemos probado, de esta manera, que todos los conjuntos $F \in \mathcal{O}_p/R$ contienen $p!$ elementos de \mathcal{O}_p . Por consiguiente, designando por $\mathcal{P}_p(E)$ al conjunto de las partes de E con p elementos, se tiene:

$$p! \text{ card}(\mathcal{P}_p(E)) = \text{card}(\mathcal{O}_p),$$

de donde se obtiene el teorema, si se tiene en cuenta I.10.4 c.q.d.

Notaciones

Al entero $\frac{n!}{p!(n-p)!}$ se le llama, a veces, «número de combinaciones de n objetos tomados de p en p », y se designa por $\binom{n}{p}$, o C_n^p . La notación $\binom{n}{p}$ tiende a imponerse.

Se verifica la relación siguiente, llamada *de Pascal*, válida para todos los enteros n, p , que verifican $1 \leq p \leq n - 1$:

$$\boxed{\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}}$$

Esta relación permite calcular los $\binom{n}{p}$ paso a paso (triángulo de Pascal).

Combinaciones con repetición

TEOREMA I.10.6

Sea E un conjunto finito con p elementos, y sea $\mathbf{N}_n = \{0, 1, 2, \dots, n\}$ el conjunto de los $n + 1$ primeros números enteros ≥ 0 .

a) El número de aplicaciones $u : E \rightarrow \mathbf{N}_n$ tales que

$$\sum_{x \in E} u(x) \leq n \quad \text{es :} \quad \binom{n+p}{p}.$$

b) El número de aplicaciones $u : E \rightarrow \mathbf{N}_n$ tales que

$$\sum_{x \in E} u(x) = n \quad \text{es :} \quad \binom{n+p-1}{p-1}.$$

Demostración. Observemos primeramente que b) resulta de a) y de la fórmula:

$$\binom{n+p}{p} = \binom{n-1+p}{p} + \binom{n+p-1}{p-1}.$$

Para demostrar a), podemos suponer que $E = \mathbf{N}_p^* = \{1, 2, \dots, p\}$.

Designemos por \mathcal{F} el conjunto de las aplicaciones $u : \mathbf{N}_p^* \rightarrow \mathbf{N}_n$ tales que $\sum_{x=1}^p u(x) \leq n$, y por \mathcal{G} el conjunto de las aplicaciones $v : \mathbf{N}_p^* \rightarrow \mathbf{N}_{n+p}^*$ estrictamente crecientes para el orden natural de \mathbf{N}_p^* y \mathbf{N}_{n+p}^* .

Vamos a definir una biyección Γ de \mathcal{F} sobre \mathcal{G} .

Para cada $u \in \mathcal{F}$, definimos una aplicación Γ_u de \mathbf{N}_p^* en \mathbf{N}_{n+p}^* por

$$x \mapsto u(1) + u(2) + \cdots + u(x) + x = \left(\sum_{y=1}^x u(y) \right) + x.$$

Para $x \geq 2$, se tiene: $\Gamma_u(x) - \Gamma_u(x-1) = u(x) + 1$. Luego Γ_u es estrictamente creciente de \mathbf{N}_p^* en \mathbf{N}_{n+p}^* . Hemos definido así una aplicación $\Gamma : u \mapsto \Gamma_u$ de \mathcal{F} en \mathcal{G} . Recíprocamente, para cada $v \in \mathcal{G}$, hacemos

$$\Delta_v(1) = v(1) - 1$$

y

$$\Delta_v(x) = v(x) - v(x-1) - 1 \quad \text{si } x \geq 2.$$

Puesto que se tiene: $\sum_{x=1}^p \Delta_v(x) = v(p) - p \leq n$, vemos que $\Delta_v \in \mathcal{F}$; observemos que $\Delta : \mathcal{G} \rightarrow \mathcal{F}$ es la aplicación $v \mapsto \Delta_v$; $\Delta \circ \Gamma$ es la aplicación idéntica de \mathcal{F} , y $\Gamma \circ \Delta$ es la aplicación idéntica de \mathcal{G} . Luego Γ es una biyección, según habíamos enunciado.

Pero evidentemente un elemento $g \in \mathcal{G}$ está unívocamente determinado si damos el conjunto $g(\mathbf{N}_p^*)$, que es una parte de \mathbf{N}_{n+p}^* con p elementos. Luego $\text{card}(\mathcal{G}) = \binom{n+p}{p}$ c.q.d.

Aplicaciones

1) El número de monomios $X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$ de grado total $\leq m$, es el número de aplicaciones u de \mathbf{N}_n^* en \mathbf{N}_m tales que $\sum_{x=1}^n u(x) \leq m$. Es, pues,

$$\binom{n+m}{n}.$$

2) El número de n -plas $(\alpha_1, \dots, \alpha_n)$ de números enteros ≥ 0 , soluciones de la ecuación $\alpha_1 + \alpha_2 + \dots + \alpha_n = m$, es

$$\binom{n+m-1}{n-1}.$$

A este número se le llama, a veces, «número de combinaciones con repetición de m objetos tomados de n en n ».

Conjuntos infinitos. Conjuntos numerables

Para comparar entre sí los diversos conjuntos infinitos, se ha elaborado la teoría de los *cardinales*. Este estudio rebasa el marco de nuestra obra. Nos contentaremos con las siguientes indicaciones:

DEFINICIÓN I.10.1

§ Dos conjuntos E, F son **equipotentes** si existe una biyección de E en F .

Si E y F son finitos, son equipotentes si, y sólo si,

$$\text{card}(E) = \text{card}(F);$$

si E es equipotente a F , y si F es equipotente a G , E es equipotente a G .

La noción de cardinal se extiende a los conjuntos infinitos.

Si E y F son infinitos y equipotentes, se demuestra que existen aplicaciones $\varphi : E \rightarrow F$ inyectivas (resp. epiyectivas) y no epiyectivas (resp. no inyectivas). Cada una de estas propiedades *caracteriza* a los conjuntos infinitos.

He aquí cómo se obtiene la noción de cardinal en los lenguajes formalizados actuales: dado un conjunto E , se demuestra que se puede *construir*, por un procedimiento *canónico*, un cierto conjunto equipotente a E , y, por lo tanto, equipotente a todos los conjuntos equipotentes a E . A este conjunto «tipo» (cuya existencia teórica está asegurada, pero del que jamás es posible hacer la construcción efectiva, puesto que esta construcción depende del axioma de la elección) se le llama **cardinal** de E , y se designa por $\text{card}(E)$. Se define la suma, el producto y la potenciación de cardinales con la ayuda de diversos procedimientos de construcción de conjuntos. Igualmente se define la *desigualdad entre cardinales*: por definición,

$$\text{card}(E) < \text{card}(F)$$

si existe una *inyección* de E en F , y si $\text{card}(E) \neq \text{card}(F)$. Se demuestra que la relación « c es un cardinal y $c < a$ » es *colectivizante* en c . Se puede, pues, hablar del conjunto de los cardinales menores que un cardinal dado. En este conjunto, la relación « $c_1 < c_2$ o $c_1 = c_2$ » es una relación de *orden total*.

En contraposición, no es posible hablar del «conjunto de todos los cardinales» (las mismas dificultades que para «el conjunto de los conjuntos»).

Las dificultades que se experimentan al definir el cardinal de un conjunto provienen del hecho de que no se pueda hablar del «conjunto de los conjuntos equipotentes de un conjunto dado». El lector interesado puede consultar [4].

Hipótesis del continuo

Por definición, el menor de los cardinales infinitos es $\text{card}(\mathbf{N})$. Se establece:

$$\text{card}(\mathbf{N}) = \aleph_0 \text{ (alef cero);}$$

a este cardinal también se le llama la *potencia del numerable*. Al cardinal de \mathbf{R} (conjunto de los números reales), se le llama *potencia del continuo*; se demuestra fácilmente que $\aleph_0 < \text{card}(\mathbf{R}) = \text{card}(\mathbf{N}^{\mathbf{N}})$; la *hipótesis del continuo* postula que no existe ningún cardinal c tal que $\text{card}(\mathbf{N}) < c < \text{card}(\mathbf{R})$. Desde 1966 (trabajos del americano Cohen), se debe considerar que esta proposición es *indecidible*. Sin embargo, la influencia de esta hipótesis sobre las matemáticas usuales es prácticamente nula.

DEFINICIÓN I.10.2

Se dice que un conjunto E es **numerable** si $\text{card}(E) = \text{card}(\mathbf{N})$, dicho de otra manera, si existe una biyección de \mathbf{N} sobre E , o también si todos los elementos de E se pueden disponer en una sucesión

$$(u_0, u_1, u_2, \dots, u_n, \dots).$$

Propiedades (que resultan de las propiedades de \mathbf{N}).

Toda parte infinita de un conjunto numerable es numerable.

Si E es infinito y si existe una *epiyección* $\varphi : \mathbf{N} \rightarrow E$, E es numerable.

Si E y F son numerables, $E \cup F$ es numerable (en efecto, si $n \mapsto u_n$ y $p \mapsto v_p$ son biyecciones de \mathbf{N} sobre E y F , la aplicación

$$f : \mathbf{N} \rightarrow E \cup F$$

tal que $f(2n) = u_n$ y $f(2n+1) = v_n$ es epiyectiva).

Por recurrencia, toda reunión finita de conjuntos numerables es numerable.

TEOREMA I.10.7

|| *El producto de dos conjuntos numerables es numerable.*

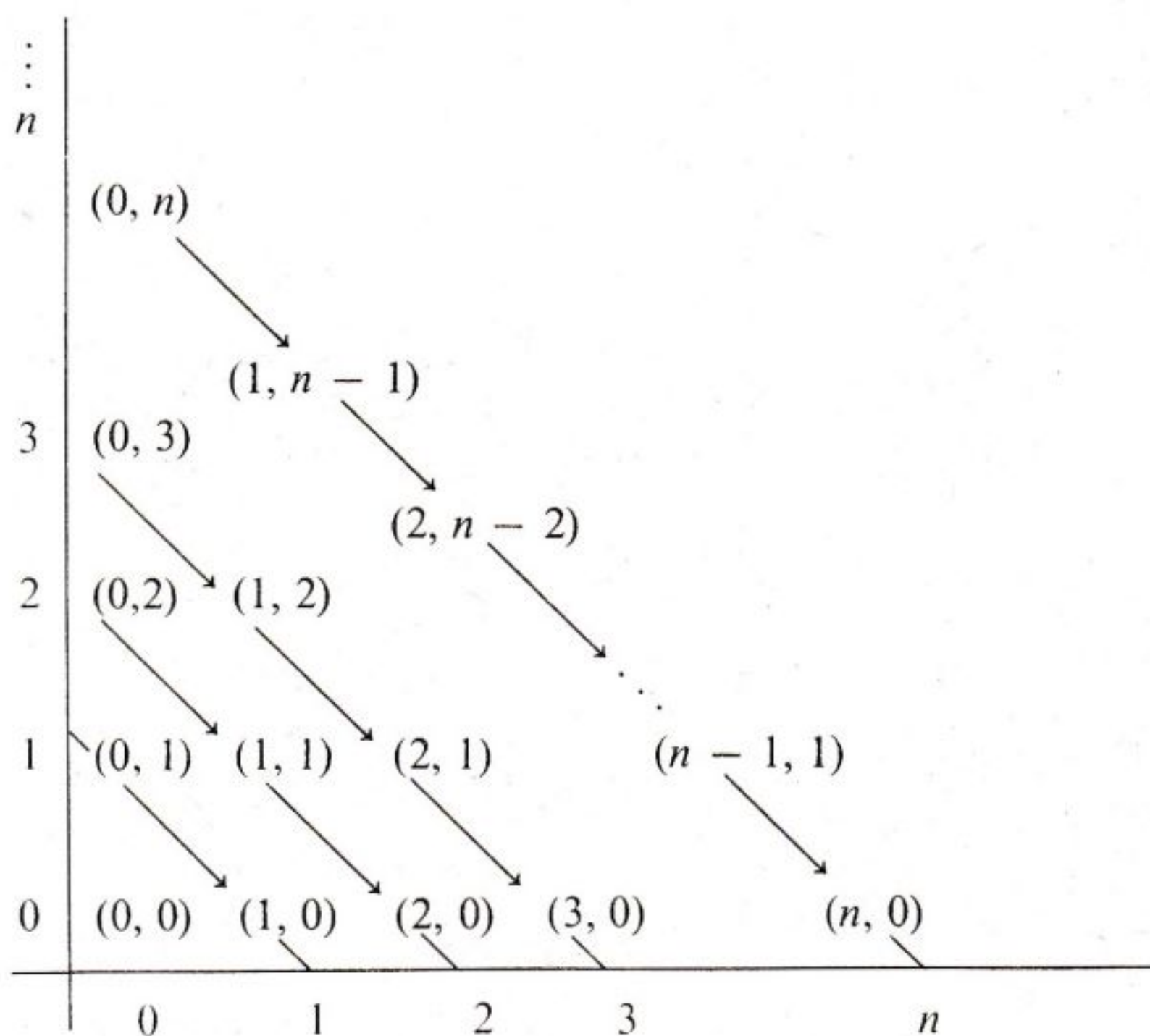
Demostración. Basta construir una biyección $\varphi : \mathbf{N} \rightarrow \mathbf{N} \times \mathbf{N}$. Se puede definir φ haciendo

$$\varphi(m) = \left(m - \frac{n(n+1)}{2}, \frac{n(n+3)}{2} - m \right)$$

para

$$\frac{n(n+1)}{2} \leq m \leq \frac{n(n+3)}{2}.$$

Ello equivale a enumerar los elementos (p, q) de $\mathbf{N} \times \mathbf{N}$ de la forma expuesta en el siguiente esquema:



Las parejas (p, q) de una diagonal seguida por las flechas son aquellas para las que $p + q = \text{Cte. c.q.d.}$

COROLARIO 1

|| El producto de una familia **finita** de conjuntos numerables es un conjunto numerable.

COROLARIO 2

|| La reunión de una familia numerable de conjuntos numerables es numerable.

Demostración. Sea $(A_i)_{i \in \mathbf{N}}$ una familia numerable de conjuntos numerables. Para cada $i \in \mathbf{N}$, sea $\theta_i : \mathbf{N} \rightarrow A_i$, $n \mapsto a_{i,n}$ una biyección de \mathbf{N} sobre A_i . Designando por $A = \bigcup_{i \in \mathbf{N}} A_i$ la reunión de los conjuntos A_i , la aplicación $\theta : \mathbf{N} \times \mathbf{N} \rightarrow A$, $(i, n) \mapsto a_{i,n}$ es una epiyección. Además, A es infinito (ya que cada A_i lo es). Por aplicación de I.10.7 y de la segunda propiedad que sigue a la definición I.10.2, se deduce que A es numerable. c.q.d.

Nota. A menudo es cómodo decir que un conjunto es numerable si es *finito o equipotente a \mathbf{N}* ; se ve fácilmente que I.10.7 y sus corolarios son también verdaderos si la palabra «numerable» se toma en esta acepción más amplia.

§ I.11 CONJUNTOS DE BASE (revisión)

Ya hemos hablado del *conjunto de los enteros naturales*, que se designa por \mathbf{N} . A lo largo de esta obra, deberemos considerar las partes de \mathbf{N} formadas por enteros inferiores o iguales a un entero dado. Emplearemos las siguientes notaciones:

$$\mathbf{N}_n = \{ p \mid p \in \mathbf{N} \text{ y } p \leq n \} \quad \mathbf{N}_n^* = \{ p \mid p \in \mathbf{N} \text{ y } 1 \leq p \leq n \} = \mathbf{N}_n \setminus \{ 0 \}$$

$$\mathbf{N}^* = \{ n \mid n \in \mathbf{N} \text{ y } n \geq 1 \} = \mathbf{N} \setminus \{ 0 \}.$$

Conjunto de los enteros relativos

Se designa por \mathbf{Z} . Recordemos simplemente que \mathbf{Z} es el cociente de $\mathbf{N} \times \mathbf{N}$ por la relación de equivalencia $(x, y) \equiv (x', y')$ definida por

$$x + y' = x' + y.$$

Conjunto de los números racionales

Se designa por \mathbf{Q} ; $\mathbf{Q} \setminus \{0\}$ se designa por \mathbf{Q}^* . En tanto que conjunto, \mathbf{Q} es el cociente de $\mathbf{Z} \times \mathbf{Z}^*$ por la relación de equivalencia $(x, y) \equiv (x', y')$ definida por $xy' = x'y$.

Se designa por \mathbf{Q}_+ el conjunto de los racionales ≥ 0 , y por \mathbf{Q}_+^* el conjunto de los racionales > 0 .

Conjunto de los números reales

Designado por \mathbf{R} , es el conjunto cociente del conjunto $\mathbf{Q}^{\mathbf{N}}$ de las sucesiones racionales, por una relación de equivalencia (cf. tomo II de la presente obra, *Análisis*).

El conjunto de los reales $\neq 0$ se designa por \mathbf{R}^* .

El conjunto de los reales ≥ 0 se designa por \mathbf{R}_+ .

El conjunto de los reales > 0 se designa por \mathbf{R}_+^* .

Conjunto de los números complejos

Designado por \mathbf{C} , es el producto $\mathbf{R} \times \mathbf{R}$. Se designa por \mathbf{C}^* al conjunto de los números complejos $\neq 0$.

Todos estos conjuntos están provistos de estructuras algebraicas que serán recordadas más adelante (para su construcción, cf. por ejemplo [12]).

Capítulo II

Leyes de composición. Grupos

§ II.1 GENERALIDADES

DEFINICIÓN II.1.1

- Una **ley de composición interna** en un conjunto E es una aplicación de $E \times E$ en E .
- Una **ley de composición externa por la izquierda** en un conjunto E , que tenga por **dominio de operadores** el conjunto Ω , es una aplicación de $\Omega \times E$ en E .
- Una **ley externa por la derecha** en E , con dominio Ω , es una aplicación de $E \times \Omega$ en E .

Con miras a las propiedades especiales por las que nos interesaremos (asociatividad, conmutatividad) es más cómodo utilizar un símbolo para designar una ley de composición (interna o externa), que utilizar la notación funcional.

Para indicar la imagen del par (x, y) por la ley \top se escribirá $x \top y$ (en vez de $\top(x, y)$) y se llamará *el compuesto* de los elementos x e y . A veces se omitirá el signo \top , y el compuesto (llamado entonces *producto*) de x e y se designará por xy .

DEFINICIÓN II.1.2

- Dotar de estructura a un conjunto E consiste en definir en E un **conjunto finito** de leyes de composición, internas o externas, sujetas a verificar

{ cierto número de condiciones, llamadas **axiomas**, de la estructura en cuestión.

Las estructuras que intervendrán en esta obra contendrán, a lo sumo, una ley externa.

Los axiomas de la estructura, independientes de la «materialización» de E , constituyen lo importante y útil; y a dos estructuras, definidas respectivamente en conjuntos E y F , se les llamará *homólogas* si:

1.º Las leyes de composición en E y F , internas o externas, coinciden en número; y las leyes externas poseen los mismos dominios de operadores.

2.º Los axiomas de las estructuras de E y F son idénticos: precisando, es posible establecer entre las leyes de E , y las de F , una correspondencia en la que las leyes que se correspondan verifiquen los mismos axiomas.

El nexo entre los diversos conjuntos provistos de estructuras homólogas dos a dos es el *morfismo*:

DEFINICIÓN II.1.3

{ Sean E, E' dos conjuntos provistos de estructuras homólogas, por lo que contienen el mismo número n de leyes internas (designadas, respectivamente \top_i, \top'_i) y el mismo número p de leyes externas (designadas, respectivamente, \perp_j, \perp'_j), ordenadas de forma que los axiomas verificados por las leyes con los mismos subíndices sean los mismos para E y E' , y que las leyes externas \perp_j y \perp'_j posean el mismo dominio de operadores Ω_j .

a) Diremos que una aplicación $f: E \rightarrow E'$ es un **morfismo** (para la estructura considerada) si:

— para todo i ($1 \leq i \leq n$), todo $x \in E$, todo $y \in E$,

$$f(x \top_i y) = f(x) \top'_i f(y),$$

— para todo j ($1 \leq j \leq p$), todo $x \in E$, todo $\lambda \in \Omega_j$,

$$f(\lambda \perp_j x) = \lambda \perp'_j f(x).$$

A un morfismo también se le llama un **homomorfismo**.

b) Un **isomorfismo** de E sobre E' es una biyección $f: E \rightarrow E'$ que es además un homomorfismo. (Se observará que su recíproca

$$f^{-1}: E' \rightarrow E$$

es un homomorfismo de E' sobre E , puesto que $z = x \top y$ equivale a

$$f(z) = f(x) \top' f(y).$$

- { c) Un **endomorfismo** de E es un morfismo de E en E .
 { d) Un **automorfismo** de E es un isomorfismo de E sobre E .

Nota. A menudo al *par* formado por el conjunto E y por una estructura algebraica definida en E se le designa sólo con la letra E , y se dice que dos conjuntos dotados de estructura E , E' son *isomorfos* si existe un isomorfismo de la estructura de E en la de E' . Se trata de un abuso de lenguaje, tolerable en tanto que no lleve a confusión. Cuando se usa este lenguaje y se desea volver al *conjunto* E , se habla del *conjunto subyacente a E* . Si E posee varias estructuras, se precisará cada vez la estructura considerada (por ejemplo, a \mathbb{R} se le puede considerar como grupo aditivo o como cuerpo; a \mathbb{C} se le puede considerar como \mathbb{R} -espacio vectorial, como \mathbb{C} -espacio vectorial o como cuerpo).

TEOREMA II.1.1

Si E , F y G son tres conjuntos dotados de estructuras homólogas, y si
 $f: E \rightarrow F$ y $g: F \rightarrow G$ son morfismos, entonces
 $g \circ f: E \rightarrow G$ es un morfismo.

Es inmediato.

Partes estables y leyes inducidas

Sea E un conjunto provisto de una estructura, definida por las leyes internas \top_1, \dots, \top_n , y por las leyes externas \perp_1, \dots, \perp_p con dominios de operadores $\Omega_1, \dots, \Omega_p$.

Se da la siguiente definición:

DEFINICIÓN II.1.4

Una parte F de E es **estable respecto de las leyes de E** si se verifican las siguientes condiciones:

a) para todo i ($1 \leq i \leq n$), todo $x \in F$ y todo $y \in F$, se tiene

$$x \top_i y \in F;$$

b) para todo j ($1 \leq j \leq p$) y todo $\lambda \in \Omega_j$, todo $x \in F$, se tiene

$$\lambda \perp_j x \in F.$$

Se puede expresar a) diciendo que, para todo i , F es estable con respecto de la ley \top_i , y b) diciendo que, para todo j , F es estable con respecto de la ley \perp_j .

Sea entonces F una parte estable de E . Designemos por \top_i'' la restricción de \top_i a $F \times F$. \top_i'' es una aplicación de $F \times F$ en E ; decir que F es estable equivale a decir que la imagen de \top_i'' está contenida en F . Si designamos por \top_i' la aplicación de $F \times F$ en F que toma los mismos valores que \top_i'' , vemos que \top_i' es una ley de composición en F .

Igualmente, sea \perp_j'' la restricción de \perp_j a $\Omega_j \times F$; \perp_j'' es una aplicación de $\Omega_j \times F$ en E que toma sus valores en F . Y la aplicación

$$\perp_j' : \Omega_j \times F \rightarrow F$$

que toma los mismos valores que \perp_j'' es una ley externa sobre F , con dominio Ω_j .

DEFINICIÓN II.1.5

*Con las anteriores notaciones, la estructura definida en F por medio de las leyes internas \perp_i' y las leyes externas \perp_j' se llama **estructura inducida** (por la de E) **en F** . A las leyes \perp_i' y \perp_j' se les llama **inducidas**.*

Notas

- 1) Esta estructura no es forzosamente homóloga a la de E , ya que las \top_i' y las \perp_j' no verifican necesariamente los mismos axiomas que las \top_i y las \perp_j .
- 2) Si esta estructura inducida es homóloga a la de E , la inyección canónica de F en E es un homomorfismo.

Estructura cociente

Sea E un conjunto provisto de una estructura definida por las leyes internas \top_1, \dots, \top_n y por las leyes externas \perp_1, \dots, \perp_p de dominios $\Omega_1, \dots, \Omega_p$. Una relación de equivalencia \mathcal{R} en E es compatible con la ley interna \top_i si las relaciones: $x \in E, y \in E, x' \in E, y' \in E, x' \mathcal{R} x$ e $y' \mathcal{R} y$ implican la relación $(x' \top_i y') \mathcal{R} (x \top_i y)$; \mathcal{R} es compatible con la ley externa \perp_j si las relaciones $\lambda \in \Omega_j, x \in E, x' \in E$ y $x' \mathcal{R} x$ implican la relación

$$(\lambda \perp_j x') \mathcal{R} (\lambda \perp_j x).$$

\mathcal{R} es compatible con la estructura de E si \mathcal{R} es compatible con todas las leyes de E .

Supongámoslo así, y designemos por \bar{E} el conjunto cociente E/\mathcal{R} , y por $p : E \rightarrow \bar{E}$ la aplicación canónica. Para $X, Y \in \bar{E}$, la clase de equivalencia $p(x \top_i y)$, en donde

$x \in X$ e $y \in Y$, depende sólo de X y de Y . Si designamos por $X \bar{\top}_i Y$ esta clase, definimos una ley interna $\bar{\top}_i$ en \bar{E} .

Análogamente, para $\lambda \in \Omega_j$ y $X \in \bar{E}$, el elemento $p(\lambda \perp_j x)$ es independiente del x elegido en X , y si lo designamos por $\lambda \perp_j X$, definimos en \bar{E} una ley externa \perp_j con dominio Ω_j .

Estas leyes se denominan *leyes cociente* de las de E por \mathcal{R} .

DEFINICIÓN II.1.6

Si \mathcal{R} designa una relación de equivalencia compatible con las leyes de E , a la estructura definida en el conjunto cociente $\bar{E} = E/\mathcal{R}$, por medio de las leyes cociente de las de E por \mathcal{R} , se le llama **estructura cociente** (de la estructura de E por \mathcal{R}).

Notas

1) La estructura cociente no es necesariamente homóloga a la de E (los axiomas de la estructura E pueden no ser satisfechos por las leyes cociente).

2) Si la estructura cociente es homóloga a la de E , la aplicación canónica $p : E \rightarrow \bar{E} = E/\mathcal{R}$ es un homomorfismo. Además, p es epiyectiva.

Recíprocamente, sean E y E' dos conjuntos provistos de estructuras homólogas, y sea $p : E \rightarrow E'$ un morfismo epiyectivo. Designamos por \mathcal{R} a la relación de equivalencia asociada a f , de forma que $(x \mathcal{R} y)$ equivalga a

$$(f(x) = f(y)) .$$

Si \top_i es una ley interna de E y \top'_i es la ley de E' correspondiente, \mathcal{R} es compatible con \top_i , pues las relaciones $(x' \mathcal{R} x)$ e $(y' \mathcal{R} y)$ implican

$$f(x') = f(x) , \quad f(y') = f(y) ,$$

de donde (puesto que f es un morfismo)

$$f(x \top_i y) = f(x) \top'_i f(y) = f(x') \top'_i f(y') = f(x' \top_i y') , \quad (x \top_i y) \mathcal{R} (x' \top_i y') .$$

Se demuestra análogamente que \mathcal{R} es compatible con las leyes externas de E . Introduzcamos la descomposición canónica de f :

$$E \xrightarrow{p} E/\mathcal{R} \xrightarrow{\bar{f}} F ,$$

en donde p es la aplicación canónica y \bar{f} la biyección canónica (cf. Cap. I, § 8).

Es inmediato que \bar{f} es un isomorfismo de E/\mathcal{R} , provisto de la estructura cociente, en F c.q.d.

Si $f: E \rightarrow F$ es un morfismo cualquiera, la imagen $f(E)$ es una parte estable de F , y se puede aplicar lo que antecede, substituyendo F por $f(E)$ provisto de la estructura inducida. Todo esto lo podemos resumir en el teorema siguiente:

TEOREMA II.1.2 (descomposición canónica de homomorfismos)

Sea $f: E \rightarrow F$ un morfismo, en donde E y F son conjuntos dotados de estructuras homólogas. Designemos por \mathcal{R} a la relación de equivalencia asociada a f ; por $p: E \rightarrow E/\mathcal{R}$ a la aplicación canónica, por

$$j: f(E) \rightarrow F$$

a la inyección canónica y por $\bar{f}: E/\mathcal{R} \rightarrow f(E)$ a la biyección canónica, de forma que $f = j \circ \bar{f} \circ p$.

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ p \downarrow & & \uparrow j \\ E/\mathcal{R} & \xrightarrow{\bar{f}} & f(E) \end{array}$$

Entonces $f(E)$ es una parte estable de F ; \mathcal{R} es compatible con las leyes de E . Si la estructura cociente de E/\mathcal{R} es homóloga a la de E , p , j y \bar{f} son morfismos, y \bar{f} es un isomorfismo cuando se dota a E/\mathcal{R} de dicha estructura y a $f(E)$ de la estructura inducida.

§ II.2 PROPIEDADES DE UNA LEY DE COMPOSICIÓN

En este §, E designa un conjunto provisto de una ley de composición interna designada por \top .

Si x_1, \dots, x_n son elementos de E , designaremos por $x_1 \top x_2 \top \dots \top x_n$, o $\bigtop_{i=1}^n x_i$, al elemento X_n de E definido recurrentemente por las relaciones de recurrencia finita:

$$X_1 = x_1, \quad X_k = X_{k-1} \top x_k \quad (k \leq n).$$

Como vamos a ver, esta notación adquiere su máximo interés cuando \top es asociativa.

DEFINICIÓN II.2.1

A la ley \top definida en E se le llama asociativa si es verdadera la siguiente relación:

$$\forall x, y, z \in E \quad (x \top y) \top z = x \top (y \top z).$$

(Según nuestros convenios, $(x \top y) \top z$ se designa también $x \top y \top z$).

TEOREMA II.2.1 (Asociatividad generalizada)

Sea \top una ley asociativa en E y sean x_1, x_2, \dots, x_n elementos de E ($n \geq 3$). Designemos por $\{J_1, \dots, J_m\}$ a una partición de \mathbf{N}_n^* tal que, para $q < r$, las relaciones $i \in J_q$ y $j \in J_r$ impliquen $i < j$. Poniendo

$$J_h = \{i_{1,h}, i_{2,h}, \dots, i_{p,h}\}, \quad \text{con} \quad i_{k,h} = k - 1 + i_{1,h} \quad (1 \leq k \leq p),$$

designemos por X_h al elemento $\bigtop_{\lambda=1}^p x_{\lambda,h}$.

Entonces se tiene $\bigtop_{i=1}^n x_i = \bigtop_{h=1}^m X_h$.

Demostración. Por recurrencia sobre n ; el teorema es verdadero para $n = 3$, por definición de asociatividad. Supongámoslo verdadero para todos los enteros menores o iguales a $n - 1$, y probemos que es verdadero para el entero n . Si $J_m = \{n\}$, se tiene

$$X_m = x_n, \quad \text{y} \quad \bigtop_{i=1}^n x_i = \left(\bigtop_{i=1}^{n-1} x_i \right) \top x_n.$$

Según la hipótesis de recurrencia,

$$\bigtop_{i=1}^{n-1} x_i = \bigtop_{h=1}^{m-1} X_h, \quad \text{luego} \quad \bigtop_{i=1}^n x_i = \left(\bigtop_{h=1}^{m-1} X_h \right) \top X_m = \bigtop_{h=1}^m X_h.$$

Si $J_m = \{p, p+1, \dots, n\}$ con $p < n$, se tiene $X_m = Y_m \top x_n$, con

$$Y_m = \bigtop_{j=p}^{n-1} x_j ;$$

luego $\bigtop_{i=1}^n x_i = \left(\bigtop_{i=1}^{n-1} x_i \right) \top x_n$; según la hipótesis de recurrencia,

$$\bigtop_{i=1}^{n-1} x_i = \left(\bigtop_{h=1}^{m-1} X_h \right) \top Y_m .$$

De esto se deduce

$$\bigtop_{i=1}^n x_i = \left\{ \left(\bigtop_{h=1}^{m-1} X_h \right) \top Y_m \right\} \top x_n ;$$

y aplicando la propiedad de la asociatividad el segundo miembro se tiene

$$\left(\bigtop_{h=1}^{m-1} X_h \right) \top (Y_m \top x_n) = \left(\bigtop_{h=1}^{m-1} X_h \right) \top X_m = \bigtop_{h=1}^m X_h . \text{ c.q.d.}$$

Este teorema significa que *para calcular el compuesto $x_1 \top x_2 \top \dots \top x_n$, los términos se pueden agrupar arbitrariamente, siempre que se conserve su orden.*

DEFINICIÓN II.2.2

Sea \top una ley en E ; un elemento e de E es **neutro por la derecha** (resp. por la izquierda) para \top si

$$\forall x \in E, x \top e = x \quad (\text{resp. } \forall x \in E, e \top x = x)$$

e es **neutro** si lo es, a la vez, por la derecha y por la izquierda.

Propiedades

— Si \mathcal{R} es una relación de equivalencia en E compatible con \top , y si e es neutro (resp. neutro por la derecha o por la izquierda), entonces la imagen $p(e)$ de e por la aplicación canónica $p: E \rightarrow E/\mathcal{R}$ es neutro (resp. neutro por la derecha o por la izquierda) para la estructura cociente.

— Para una ley dada, E posee, a lo sumo, un elemento neutro. En efecto, si e, e' designan dos elementos neutros, se tendrá

$$e \top e' = e = e' ,$$

DEFINICIÓN II.2.3

Un elemento $a \in E$ es **regular por la derecha** (para la ley interna \top en E) si la aplicación $x \mapsto x \top a$ es inyectiva, en otras palabras, si $x \top a = y \top a$ implica $x = y$; es **regular por la izquierda** si la aplicación $x \mapsto a \top x$ es inyectiva; a es **regular** si es regular por la derecha y por la izquierda.

La relación: $((x \top a = y \top a) \Rightarrow (x = y))$ significa que las igualdades «que contienen a a como factor» se pueden «simplificar por a ».

Las nociones de neutro por la derecha o por la izquierda, de regular por la derecha o por la izquierda, coinciden si la ley \top es *conmutativa*.

DEFINICIÓN II.2.4

La ley interna \top en E es **conmutativa** si, para todo $x \in E$ y todo $y \in E$, se tiene:

$$x \top y = y \top x .$$

● Las propiedades de *asociatividad* y *conmutatividad* se conservan siempre por paso a la estructura cociente o a la estructura inducida. (Tanto si dichas estructuras son homólogas a la estructura dada, como si no lo son.) Esta observación es importante en la práctica.

Ejemplo

Para probar la asociatividad y la conmutatividad de la suma y de la multiplicación de \mathbf{Q} , es suficiente establecer estas propiedades en $\mathbf{Z} \times \mathbf{Z}^*$ (cf. T. III.6.6).

Si \top es, a la vez, asociativa y conmutativa, se tiene el siguiente teorema:

TEOREMA II.2.2

Sea \top una ley interna en E asociativa y conmutativa, y sean

$$x_1, x_2, \dots, x_n$$

elementos de E . Para toda permutación σ del conjunto de los enteros $\{1, 2, \dots, n\}$ se tiene

$$\prod_{i=1}^n x_i = \prod_{i=1}^n x_{\sigma(i)}.$$

Al resultado se le podrá designar entonces por $\prod_{1 \leq i \leq n} x_i$ (sin precisar el orden de las x_i).

Demostración. Por recurrencia sobre n ; para $n = 2$ el resultado se sigue de las definiciones.

Suponiéndolo verdadero para los enteros inferiores o iguales a $n - 1$, demos-trémoslo para el entero n :

— Si $\sigma(n) = n$, podemos escribir:

$$\prod_{i=1}^n x_{\sigma(i)} = \left(\prod_{i=1}^{n-1} x_{\sigma(i)} \right) \top x_n;$$

por la hipótesis de recurrencia, y teniendo en cuenta que la restricción de σ a $\{1, 2, \dots, n - 1\}$ es una permutación de estos enteros, se tiene

$$\prod_{i=1}^{n-1} x_{\sigma(i)} = \prod_{i=1}^{n-1} x_i \quad \text{de donde} \quad \prod_{i=1}^n x_{\sigma(i)} = \prod_{i=1}^n x_i.$$

— Si $\sigma(n) = p < n$, sea τ la permutación de $\{1, 2, \dots, n\}$ definida por

$$\tau(n) = \sigma(n-1), \quad \tau(\sigma(n-1)) = n, \quad \text{y} \quad \tau(k) = k \quad \text{si} \quad k \neq n \quad \text{y} \quad k \neq \sigma(n-1);$$

(τ cambia $\sigma(n-1)$ con n y deja p fijo).

Según la hipótesis de recurrencia se puede escribir

$$\prod_{i=1}^{n-1} x_{\sigma(i)} = \prod_{i=1}^{n-1} x_{\tau(\sigma(i))} = \left(\prod_{i=1}^{n-2} x_{\tau \circ \sigma(i)} \right) \top x_n$$

luego, por asociatividad:

$$\prod_{i=1}^n x_{\sigma(i)} = \left(\left(\prod_{i=1}^{n-2} x_{\tau \circ \sigma(i)} \right) \top x_n \right) \top x_{\sigma(n)} = \left(\prod_{i=1}^{n-2} x_{\tau \circ \sigma(i)} \right) \top (x_n \top x_p)$$

y, por conmutatividad, esto es igual a

$$\left(\prod_{i=1}^{n-2} x_{\tau \circ \sigma(i)} \right) \top (x_p \top x_n) = \left(\left(\prod_{i=1}^{n-2} x_{\tau \circ \sigma(i)} \right) \top x_p \right) \top x_n.$$

Volvemos pues al caso anterior, c.q.d.

§ II.3 AXIOMAS DE LA ESTRUCTURA DE GRUPO. EJEMPLOS DE GRUPOS. HOMOMORFISMOS

DEFINICIÓN II.3.1

Sea G un conjunto, provisto de una ley de composición

$$(x, y) \mapsto x \cdot y.$$

Se dice que esta ley determina en G una estructura de **grupo** (o, más brevemente, que G es un **grupo**) si se verifican los siguientes axiomas:

- (G₁) la ley es asociativa,
- (G₂) en G existe un elemento neutro,
- (G₃) para todo elemento x de G , existe un elemento x' de G tal que $x \cdot x' = x' \cdot x = e$, en donde e designa el elemento neutro.

Las propiedades siguientes son inmediatas:

- Un grupo es *no vacío* (puesto que contiene el elemento neutro).
- El elemento neutro es único (cf. § 2).
- Sean e el elemento neutro y x un elemento cualquiera de G . Existe un único elemento x' tal que $x \cdot x' = x' \cdot x = e$. En efecto si $x \cdot x'' = x'' \cdot x = e$, se tiene por asociatividad

$$\begin{aligned} x' \cdot x \cdot x'' &= (x' \cdot x) x'' = e \cdot x'' = x'' \\ &= x' \cdot (x \cdot x'') = x' \cdot e = x'. \end{aligned}$$

A este elemento se le llama el *simétrico* de x . Cuando a la ley del grupo se le llama *multiplicación*, al elemento simétrico de x también se le llama el *inverso* de x , y se le designa por x^{-1} ; y el elemento neutro se designa por e , o por 1 .

- Todos los elementos de un grupo son regulares.

En efecto, sea x un elemento de G . Si $x \cdot x' = x \cdot x''$, se deduce por asociatividad:

$$x^{-1} \cdot (x \cdot x') = x^{-1} \cdot (x \cdot x'') = (x^{-1} \cdot x) \cdot x' = (x^{-1} \cdot x) \cdot x'' = e \cdot x' = e \cdot x'' = x' = x''.$$

Igualmente, si $x' \cdot x = x'' \cdot x$, se tiene $x' = x''$.

DEFINICIÓN II.3.2

$\left\{ \begin{array}{l} \text{Sea } G \text{ un grupo; se dice que } G \text{ es } \mathbf{abeliano} \text{ (o conmutativo) si la ley de} \\ G \text{ es conmutativa.} \end{array} \right.$

Habitualmente la ley de un grupo abeliano se designa *aditivamente*, es decir, por medio del signo de adición $+$. Al compuesto $x + y$ de x e y se le llama entonces *suma* de x e y .

El elemento neutro se designa por 0 , y al elemento simétrico de un elemento x se le llama *opuesto de x* y se designa por $-x$; al elemento $x + (-y)$ se le designa por $x - y$.

En un grupo abeliano se verifican las siguientes reglas de cálculo:

$$x - (y + z) = (x - y) - z = x - y - z$$

$$x - (y - z) = x - y + z, \text{ etc. (regla de los signos).}$$

Notas

1) Los axiomas (G_1) , (G_2) y (G_3) pueden ser debilitados de forma considerable (cf. ejercicios).

2) En general, un grupo cualquiera se designa multiplicativamente, y un grupo abeliano, aditivamente; sin embargo, cuando en una estructura intervienen varias leyes, nos vemos obligados a designar multiplicativamente ciertos grupos abelianos; ejemplo: \mathbf{Q}^* (ejemplo 3).

El empleo de un mismo símbolo para designar la ley interna de grupos distintos (que, en principio, podría prestarse a confusión), en la práctica no resulta molesta, y tampoco introduce equívocos en los textos.

Ejemplos

1) Sean E un conjunto no vacío, y \mathfrak{S}_E el conjunto de las biyecciones de E en E . Dotemos a \mathfrak{S}_E de la ley de composición $(f, g) \mapsto f \circ g$ (composición de aplicaciones). Entonces \mathfrak{S}_E se convierte en un grupo; el elemento neutro es la aplicación identidad $I : E \rightarrow E$, tal que $I(x) = x$ para todo x . El inverso de $f \in \mathfrak{S}_E$ es la biyección recíproca de f (que hemos designado por f^{-1} en el § 8 del capítulo I). A este grupo se le llama el *grupo de las permutaciones* de E ; si E tiene, por lo menos, 3 elementos, \mathfrak{S}_E no es abeliano (ver § 7).

Cuando E es el conjunto $\{1, 2, \dots, n\}$ de los n primeros enteros > 0 , al grupo \mathfrak{S}_E se le designa por \mathfrak{S}_n y se le llama *grupo simétrico* de orden n ; entonces se tiene: $\text{card}(\mathfrak{S}_n) = n!$ (ver § I. 10).

2) El conjunto \mathbf{Z} de los números enteros, provisto de la suma, forma un grupo abeliano, cuyo elemento neutro es 0.

3) El conjunto \mathbf{Q}^* de los racionales $\neq 0$, provisto de la multiplicación, forma un grupo abeliano, cuyo elemento neutro es 1. También lo es el subconjunto del anterior \mathbf{Q}_+^* , formado por los racionales > 0 .

4) Sean G un grupo y E un conjunto cualquiera. Dotemos al conjunto $\mathcal{F}(E, G)$ de las aplicaciones de E en G de la ley siguiente:

Si f y g son aplicaciones de E en G , $f \cdot g$ es la aplicación

$$x \mapsto f(x) \cdot g(x)$$

de E en G .

Entonces $\mathcal{F}(E, G)$ es un grupo, cuyo elemento neutro es la aplicación *constante* $x \mapsto e$ (en donde e es el elemento neutro de G). La aplicación inversa de $f: E \rightarrow G$ es la aplicación $x \mapsto (f(x))^{-1}$.

A este grupo se le designa a veces por G^E : esta notación se precisará al estudiar los grupos producto.

DEFINICIÓN II.3.3 (Caso particular de II.1.3)

Sean G_1 y G_2 dos grupos (designados multiplicativamente). Un **homomorfismo** de G_1 en G_2 es una aplicación $f: G_1 \rightarrow G_2$ tal que, para todo $x \in G_1$ y todo $y \in G_1$, se verifique

$$f(x \cdot y) = f(x) \cdot f(y).$$

Si, además, f es biyectiva, f es un **isomorfismo**.

Finalmente un **automorfismo** del grupo G es un isomorfismo de G en sí mismo

Propiedades

— Un homomorfismo transforma el elemento neutro e_1 de G_1 en el elemento neutro e_2 de G_2 : en efecto, se tiene

$$f(e_1 \cdot e_1) = f(e_1) = f(e_1) \cdot f(e_1) = f(e_1) \cdot e_2,$$

de donde $f(e_1) = e_2$ puesto que $f(e_1)$ es regular.

— Si $x \in G_1$, se tiene que $f(x^{-1}) = (f(x))^{-1}$, pues

$$f(x^{-1}) \cdot f(x) = f(x^{-1} \cdot x) = f(e_1) = e_2.$$

Ejemplos

1) Si a es un elemento del grupo G , la aplicación $x \mapsto a^{-1}xa$ de G en G es un homomorfismo, ya que:

$$(a^{-1}xa)(a^{-1}ya) = a^{-1}x(aa^{-1})ya = a^{-1}(xy)a.$$

Este homomorfismo es, en realidad, un automorfismo, puesto que es biyectivo y admite por recíproco el homomorfismo $x \mapsto axa^{-1}$.

A los automorfismos de esta forma se les llama los **automorfismos internos** de G .

2) Designamos por $f: E \rightarrow F$ a una biyección del conjunto E en el conjunto F . La aplicación $f^*: \mathfrak{S}_F \rightarrow \mathfrak{S}_E$, definida, para todo $\varphi \in \mathfrak{S}_F$, por

$$f^*(\varphi) = f^{-1} \circ \varphi \circ f,$$

es un isomorfismo de grupos.

En particular, si E es finito, se ve que \mathfrak{S}_E es isomorfo a \mathfrak{S}_n (con $n = \text{card}(E)$), de $n!$ formas distintas, y no existe ningún isomorfismo privilegiado de \mathfrak{S}_E en \mathfrak{S}_n . La elección de uno de estos isomorfismos equivale a dar una relación de orden total en E .

3) Sea G un grupo, designado multiplicativamente, con elemento neutro e . Para todo $n \in \mathbf{Z}$, definimos x^n por:

$$x^0 = e, \quad x^n = x^{n-1} \cdot x \quad \text{para } n \geq 1; \quad \text{y si } n < 0.$$

$$x^n = (x^{-1})^{-n}.$$

La aplicación $n \mapsto x^n$ es entonces un homomorfismo de \mathbf{Z} en G , pues la fórmula $x^{n+p} = x^n \cdot x^p$ es válida para $n, p \in \mathbf{Z}$.

Cuando G se escribe aditivamente, a esta aplicación se le designa por $n \mapsto n \cdot x$, y la fórmula anterior se convierte en:

$$(n + p) \cdot x = n \cdot x + p \cdot x \quad (n, p \in \mathbf{Z}).$$

4) Si $f: G_1 \rightarrow G_2$ es un homomorfismo de grupos, y E es un conjunto, definimos la aplicación ${}^t f: G_1^E \rightarrow G_2^E$, por $({}^t f)(\varphi) = f \circ \varphi$ para toda aplicación $\varphi: E \rightarrow G_1$; ${}^t f$ es un homomorfismo de grupos. (Para la definición de G^E , ver el ejemplo 4 precedente.)

5) Sean G un grupo, E y F dos conjuntos y $f: E \rightarrow F$ una aplicación. La aplicación $f^*: G^F \rightarrow G^E$, tal que $f^*(\varphi) = \varphi \circ f$ (φ , aplicación de F en G) es un homomorfismo de grupos.

6) Si G_1 y G_2 son abelianos, el conjunto designado por $\text{Hom}(G_1, G_2)$, de los homomorfismos de G_1 en G_2 , es una parte estable de $G_2^{G_1}$; y, provisto de la ley inducida, es un grupo abeliano.

§ II.4 SUBGRUPO. GRUPO ENGENDRADO. GRUPO PRODUCTO

En este párrafo los grupos se escribirán multiplicativamente.

DEFINICIÓN II.4.1

Sean G un grupo y H una parte de G ; H es un **subgrupo** de G si H es una parte estable de G , y si, con la ley inducida por la de G , H es un grupo.

Sea H un subgrupo de G ; designemos por e al elemento neutro de G y por e' al de H . Se tiene que $e \cdot e' = e' = e' \cdot e$, de donde $e = e'$ puesto que e' es regular en G . Luego e es siempre elemento de H , y es el elemento neutro de H .

TEOREMA II.4.1

Sea G un grupo. Para que una parte H de G sea un subgrupo de G , es necesario y suficiente que se verifiquen las condiciones siguientes:

- a) H es no vacío;
- b) H es estable;
- c) para todo $x \in H$, $x^{-1} \in H$.

Demostración

1) Si H es un subgrupo, a) y b) se verifican. Sea $x \in H$ y x^{-1} el simétrico de x en G , x'^{-1} el simétrico de x en H . Se tiene

$$x \cdot x^{-1} = x \cdot x'^{-1} = e$$

puesto que el neutro de H es el mismo que el de G , de donde $x^{-1} = x'^{-1}$ ya que x es regular en G . Luego c) también se verifica.

2) Si se verifican a), b), c), demostraremos que la ley inducida por G en H es una ley de grupo: puesto que $H \neq \emptyset$, existe un $a \in H$. Según c), $a^{-1} \in H$, de donde, por b), $a \cdot a^{-1} = e \in H$. Luego H posee el elemento neutro e . Aplicando c), de nuevo, vemos que H es un grupo. c.q.d.

Nota. b) y c) equivalen a la condición única:

$$(x, y \in H) \Rightarrow (xy^{-1} \in H).$$

Propiedades de los subgrupos (su verificación no presenta ninguna dificultad)

- $\{e\}$ y G son subgrupos de G . A los otros subgrupos se les llama subgrupos *propios*. Todo subgrupo de G contiene a $\{e\}$.
- Si H es un subgrupo de G , y K es un subgrupo de H , K es un subgrupo de G .
- Si $(H_i)_{i \in I}$ es una familia de subgrupos de G , $\bigcap_{i \in I} H_i$ es un subgrupo de G .
- Sean G_1, G_2 dos grupos, y $f: G_1 \rightarrow G_2$ un homomorfismo. Para todo subgrupo H_1 de G_1 , la imagen directa $f(H_1)$ es un subgrupo de G_2 ; para todo subgrupo H_2 de G_2 , la imagen recíproca $f^{-1}(H_2)$ es un subgrupo de G_1 . En particular, la imagen recíproca $f^{-1}(e_2)$, en donde e_2 es el elemento neutro de G_2 , es un subgrupo de G_1 . Son notables los siguientes subgrupos:

DEFINICIÓN II.4.2

Sean G_1 y G_2 dos grupos con elementos neutros e_1 y e_2 . Si

$$f: G_1 \rightarrow G_2$$

es un homomorfismo, al subgrupo $f^{-1}(e_2)$ de G_1 se le llama **núcleo de f** y se designa por $\text{Ker } f$; al subgrupo de G_2 formado por $f(G_1)$ se le llama **imagen de f** y se designa por $\text{Im } f$.

Conservemos las notaciones de la definición II.4.2; fijemos $x \in G_1$, y hagamos

$$y = f(x), \quad N = \text{Ker } f.$$

La relación $f(x') = y$ implica:

$$(f(x))^{-1} f(x') = e_2 = f(x^{-1}) f(x') = f(x^{-1} x'),$$

de donde $x^{-1}x' \in N$. Se puede ver también que $f(x') = y$ implica:

$$x' x^{-1} \in N.$$

Luego la imagen recíproca $f^{-1}(y)$ es igual al conjunto de elementos $(x.v)_{v \in N}$, y también al conjunto de elementos $(v.x)_{v \in N}$. En particular:

TEOREMA II.4.2

|| Para que el homomorfismo de grupos $f: G_1 \rightarrow G_2$ sea inyectivo es necesario y suficiente que su núcleo se reduzca a $\{e\}$.

Consideremos de nuevo el homomorfismo $f: G_1 \rightarrow G_2$, y sea H_2 un subgrupo de G_2 ; $f^{-1}(H_2)$ es un subgrupo de G_1 , que contiene a N . Recíprocamente, para todo subgrupo H_1 de G_1 que contenga a N , $f(H_1)$ es un subgrupo de G_2 , y según el estudio precedente, se tiene que $f^{-1}(f(H_1)) = H_1$. Si, además, f es epiyectiva, se tiene también: $f(f^{-1}(H_1)) = H_1$.

De donde:

Si $f: G_1 \rightarrow G_2$ es un homomorfismo de grupos epiyectivo, la aplicación $H_2 \mapsto f^{-1}(H_2)$ es una biyección del conjunto de los subgrupos de G_2 en el conjunto de los subgrupos de G_1 que contienen a N , cuya biyección recíproca es la aplicación $H_1 \mapsto f(H_1)$.

DEFINICIÓN II.4.3

Sean G un grupo y A una parte de G . El subgrupo engendrado por A en G es la intersección del conjunto de los subgrupos de G que contienen a A . Lo designaremos por $\text{gr}(A)$, y por $\text{gr}(x)$ si A se reduce al elemento x .

Ordenemos el conjunto \mathcal{G} de los subgrupos de G por inclusión: se puede decir también que $\text{gr}(A)$ es el ínfimo del conjunto \mathcal{G}_A de los subgrupos de G que contienen a A ; $\text{gr}(A)$ es también el ínfimo de \mathcal{G}_A en $\mathcal{P}(G)$, ordenado por inclusión.

Si A es no vacío, $\text{gr}(A)$ contiene todos los elementos de G de la forma

$$\prod_{i=1}^n x_i^{\alpha_i},$$

en donde $x_i \in A$ y $\alpha_i = \pm 1$. Recíprocamente, el conjunto de estos elementos forma un subgrupo de G , luego $\text{gr}(A)$ es igual a este conjunto.

Convendremos que $\text{gr}(\emptyset) = \{e\}$.

Como caso particular, tomemos una familia $(G_i)_{i \in I}$ de subgrupos de G . El grupo engendrado por $\bigcup_{i \in I} G_i$ es el menor elemento (en \mathcal{G} , ordenado por inclusión) del conjunto de los subgrupos que contienen a todos los G_i . Dicho de otra manera, la familia $(G_i)_{i \in I}$ posee un supremo en \mathcal{G} .

Pero este supremo no coincide con el supremo de $(G_i)_{i \in I}$ en $\mathcal{P}(G)$ que es $\bigcup_{i \in I} G_i$, ni tan siquiera cuando I es finito (cf. ejercicios) ya que $\bigcup_{i \in I} G_i$, en general, no es un subgrupo de G .

Ejemplos

1) Subgrupos de \mathbf{Z}

Sea a un entero prefijado y sea $a\mathbf{Z}$ el conjunto de los enteros relativos de la forma $n \cdot a$ ($n \in \mathbf{Z}$); $a\mathbf{Z}$ es un subgrupo de \mathbf{Z} . Recíprocamente, sea H un subgrupo

de \mathbf{Z} y veamos que existe un $a \in \mathbf{N}$ tal que $H = a\mathbf{Z}$; si $H = \{0\}$ es evidente. Si no, H contiene elementos $\neq 0$, por lo tanto elementos > 0 ya que $(x \in H)$ implica $((-x) \in H)$. Sea a el primera de los elementos > 0 de H ; H contiene a $a\mathbf{Z}$ (cf. ejemplo 2); además, si $m \in H$, efectuamos la división euclídea de m por a :

$$m = aq + r, \quad 0 \leq r < a, \quad q \in \mathbf{Z}.$$

Puesto que $m \in H$, se tiene: $r = m - aq \in H$, de donde $r = 0$ ya que a es el primer elemento de H que es > 0 ; esto demuestra que $H = a\mathbf{Z}$.

2) Orden de un elemento en un grupo

Sea G un grupo y x un elemento de G . La imagen del homomorfismo $n \mapsto x^n$ de \mathbf{Z} en G es un subgrupo de G ; evidentemente esta imagen es $\text{gr}(x)$. El núcleo de este homomorfismo es de la forma $\omega\mathbf{Z}$, en donde ω es un entero. Si este entero es nulo, $n \mapsto x^n$ es una inyección, y $\text{gr}(x)$ es isomorfo a \mathbf{Z} . En este caso se dice que x es de orden infinito.

Si $\omega > 0$, ω posee la siguiente propiedad, que es característica:

$$x^\omega = e \quad (e \text{ neutro de } G).$$

y todo entero n tal que $x^n = e$ es múltiplo de ω .

En este caso, $\text{gr}(x)$ es finito y está formado por los elementos: $e, x, x^2, \dots, x^{\omega-1}$, cuyo número es ω . Se dice entonces que x es de orden ω .

En virtud de la importancia de los ejemplos 1) y 2), daremos las siguientes definiciones:

DEFINICIÓN II.4.4

El elemento x del grupo G es **de orden infinito** si el homomorfismo

$$n \mapsto x^n$$

de \mathbf{Z} en G es inyectivo, es decir, si los $(x^n)_{n \in \mathbf{Z}}$ son distintos, dos a dos.
En caso contrario, el elemento x es **de orden finito**. Su **orden** es entonces el entero $\omega > 0$ tal que el núcleo del homomorfismo $n \mapsto x^n$ de \mathbf{Z} en G es $\omega\mathbf{Z}$; ω se puede definir también como el menor de los enteros $m \geq 1$ tales que $x^m = e$.

Más adelante utilizaremos el siguiente resultado:

PROPOSICIÓN II.4.3

En un grupo G , sean x e y dos elementos de órdenes respectivos α y β . Si α y β son **primos entre sí**, y si $xy = yx$, entonces el orden γ de xy es: $\gamma = \alpha\beta$.

Demostración

a) De la hipótesis $xy = yx$, se deduce

$$\forall m \in \mathbf{Z} \quad (xy)^m = x^m y^m.$$

luego

$$(xy)^\gamma = x^\gamma y^\gamma = e,$$

por lo que γ es múltiplo del orden de xy .

b) Recíprocamente, sea $m \in \mathbf{Z}$ tal que $(xy)^m = e$. Se deduce:

$$(xy)^{m\beta} = e = x^{m\beta} y^{m\beta} = x^{m\beta} \cdot e = x^{m\beta},$$

luego $m\beta$ es múltiplo del orden α de x . Puesto que α y β son primos entre sí, m es múltiplo de α . Se demuestra análogamente que m es múltiplo de β : y dado que α y β son primos entre sí, m es múltiplo de γ . c.q.d.

3) Sea E un conjunto no vacío, y G un subgrupo de \mathfrak{S}_E (a G se le denomina *grupo de permutaciones de E*). Para toda parte no vacía A de E , el conjunto G_A de las $f \in G$ tales que $f(A) = A$, es un subgrupo de G . Para todo $a \in E$, el conjunto $G_{\{a\}}$ es un subgrupo de G , designado a veces por G_a , y denominado *subgrupo de isotropía de a* .

Para toda parte no vacía A de E , el conjunto de las $f \in G$ tales que, para todo $a \in A$, $f(a) = a$, es un subgrupo H_A de G_A : es la intersección de los $(G_a)_{a \in A}$.

El lector deberá tener cuidado en no confundir H_A y G_A : A es *invariante globalmente* para todo elemento de G_A , e *invariante punto a punto* para todo elemento de H_A (cf. § 8).

Hagamos $B = E \setminus A$. La aplicación $\rho : H_A \rightarrow \mathfrak{S}_B$, que, a toda $f \in H_A$, le asocia su restricción a B , es un isomorfismo de H_A en \mathfrak{S}_B .

En particular, tomando $E = \mathbf{N}_n^*$ y $B = \mathbf{N}_p^*$, con $p < n$, se ve que \mathfrak{S}_p se identifica canónicamente con el subgrupo de las permutaciones de E que dejan fijos cada uno de los enteros $p+1, p+2, \dots, n$.

Grupo producto**TEOREMA II.4.4**

Sea $(G_i)_{i \in I}$ una familia de grupos. Dotemos al conjunto producto $G = \prod_{i \in I} G_i$ de la siguiente ley interna:

$$\text{si } x = (x_i)_{i \in I} \text{ y } y = (y_i)_{i \in I}, \quad x \cdot y = (x_i \cdot y_i)_{i \in I}.$$

Entonces G es un grupo.

Demostración. Si e_i es elemento neutro de G_i , $e = (e_i)_{i \in I}$ es elemento neutro de G . La asociatividad de G resulta de la asociatividad en cada G_i . El inverso de $x = (x_i)_{i \in I}$ es $(x_i^{-1})_{i \in I}$. c.q.d.

DEFINICIÓN II.4.5

$\left\{ \begin{array}{l} \text{Al grupo } G = \prod_{i \in I} G_i \text{ así definido se le llama } \textbf{grupo producto de la} \\ \text{familia } (G_i)_{i \in I}. \end{array} \right.$

Sea $p_i : G \rightarrow G_i$ la i -ésima proyección (cf. Cap. I) tal que, si

$$x = (x_i)_{i \in I}, \quad p_i(x) = x_i;$$

p_i es un homomorfismo epiyectivo cuyo núcleo es isomorfo al grupo $\prod_{\substack{j \in I \\ j \neq i}} G_j$. Sea

$q_j : G_j \rightarrow G$ la inyección tal que

$$q_j(x_j) = (y_i)_{i \in I}, \quad \text{con } y_j = x_j \text{ y } y_i = e_i \text{ si } i \neq j;$$

q_j es un homomorfismo, que con su ayuda es posible identificar G_j con un subgrupo de $\prod_{i \in I} G_i$. Se tiene, evidentemente, $p_i \circ p_i = p_i$, $p_i \circ p_j = \text{Cte} = e_i$ si $i \neq j$, $p_i \circ q_i =$ aplicación idéntica de G_i , y $p_i \circ q_j = e_i$ si $i \neq j$. Finalmente, sea H un grupo y $f : H \rightarrow G$ un homomorfismo. Los $p_i \circ f$ son homomorfismos de H en G_i , y para todo $x \in H$, se tiene

$$f(x) = (p_i \circ f(x))_{i \in I}.$$

Vemos, pues, que el conocimiento del homomorfismo f es equivalente al de la familia de homomorfismos $p_i \circ f : H \rightarrow G_i$ ($i \in I$).

En particular, cuando $G_i = G$ para todo $i \in I$, el grupo producto $\prod_{i \in I} G_i$ se identifica con el grupo de las aplicaciones de I en G definido en el § 3, y se le designa también por G^I .

§ II.5 GRUPO COCIENTE EN EL CASO ABELIANO

Los grupos abelianos forman una clase muy estable de grupos: un subgrupo de un grupo abeliano es abeliano, un producto de grupos abelianos es abeliano. Para los grupos abelianos vamos a estudiar especialmente el *cociente* de un grupo por un subgrupo, que es otra operación respecto de la cual la clase de los grupos abelianos es estable.

Designemos por G un grupo abeliano, que escribiremos aditivamente, y por H un subgrupo de G . Introducimos una relación binaria entre los elementos x, y de G , designada por

$$x \equiv y (H),$$

con la siguiente definición:

$$(1) \quad (x \equiv y (H)) \Leftrightarrow (x - y \in H).$$

Demostraremos que *esta relación es una relación de equivalencia en G , compatible con la ley de grupo de G .*

Se tiene $x - x = 0 \in H$, de donde $x \equiv x (H)$ para todo x .

Para todos $x, y \in G : x - y \in H \Rightarrow y - x = -(x - y) \in H$ (puesto que H es un subgrupo), de donde

$$x \equiv y (H) \Rightarrow y \equiv x (H);$$

y para toda terna $x, y, z \in G$:

$$(x - y \in H \text{ y } y - z \in H) \Rightarrow (x - z = x - y + (y - z) \in H)$$

puesto que H es estable. De donde se sigue la transitividad, y (1) es una relación de equivalencia. Si $x' \equiv x (H)$ e $y' \equiv y (H)$, vemos que $x' + y' \equiv x + y (H)$; se tiene: $x' + y' - (x + y) = x' - x + y' - y$ ya que G es abeliano, y

$$x' - x \in H, \quad y' - y \in H,$$

luego $x' - x + y' - y \in H$ puesto que H es estable. Luego (1) es compatible con la adición de G .

Recíprocamente, sea \mathcal{R} una relación de equivalencia en G , compatible con la adición de G , y designemos por H la clase de equivalencia de 0. Vemos que H es un subgrupo y que las relaciones $x \mathcal{R} y$ y $x \equiv y (H)$ son equivalentes:

Sea x un elemento de H : se tiene $x \mathcal{R} 0$, de donde, puesto que \mathcal{R} es compatible con la suma:

$$(x + (-x)) \mathcal{R} (-x), \quad 0 \mathcal{R} (-x), \quad (-x) \mathcal{R} 0, \quad \text{luego: } -x \in H.$$

Si x, y son dos elementos de H , se tiene: $x \mathcal{R} 0, y \mathcal{R} 0$, de donde por la compatibilidad de $\mathcal{R} : (x + y) \mathcal{R} (0 + 0)$, es decir, $x + y \in H$, y queda así demostrado que H es un subgrupo.

Sean finalmente x e y dos elementos de G . Puesto que \mathcal{R} es compatible con la adición, las relaciones $x \mathcal{R} y$ y $(x - y) \mathcal{R} (y - y)$ son equivalentes, de ahí la última de nuestras afirmaciones. En resumen hemos demostrado:

TEOREMA II.5.1

|| Sea G un grupo abeliano y H un subgrupo. La relación $x \equiv y (H)$ es una relación de equivalencia en G compatible con la adición. Recíprocamente,

|| *toda relación de equivalencia compatible con la adición en G es de esta forma y H es la clase del 0.*

Notación. Si H es un subgrupo del grupo abeliano G , designaremos por G/H el conjunto cociente de G por la relación de equivalencia:

$$x \equiv y (H).$$

Designaremos por $p : G \rightarrow G/H$ a la aplicación canónica.

Según las generalidades del § 1, se puede dotar a G/H de una ley cociente (que designaremos también por el símbolo $+$), definida por las relaciones siguientes:

$$(2) \quad \text{Si } X \in G/H \text{ y } Y \in G/H, \quad x \in X \text{ y } y \in Y, \\ \text{se tiene} \quad X + Y = p(x + y)$$

(según el § 1, $p(x + y)$ sólo depende de X e Y).

Se tiene, entonces, el siguiente teorema:

TEOREMA II.5.2

|| *Con las notaciones anteriores, la ley cociente definida en G/H dota a este conjunto de una estructura de **grupo abeliano**.*

A este grupo se le designa por G/H y se le llama **grupo cociente de G por H** .

Demostración.

— Asociatividad. Sean $X \in G/H$, $Y \in G/H$, $Z \in G/H$, $x \in X$, $y \in Y$, $z \in Z$. Se tiene

$$\begin{aligned} (X + Y) + Z &= p(x + y) + p(z) = p((x + y) + z) \\ &= p(x + (y + z)) = p(x) + p(y + z) \\ &= X + (Y + Z), \end{aligned}$$

en virtud de la asociatividad de la suma en G .

— Conmutatividad. Resulta igualmente de la conmutatividad de la suma en G .

— Elemento neutro. Sea $O = p(o)$, en donde o es el elemento neutro de G . Si $X \in G/H$ y $x \in X$, se tiene $X + O = p(x + o) = p(x) = X$, luego O es neutro en G/H .

— Elemento opuesto. Sean $X \in G/H$ y $x \in X$. Se tiene

$$X + p(-x) = p(x) + p(-x) = p(x + (-x)) = p(o) = O,$$

luego $p(-x) = -X$. c.q.d.

Así vemos que la propiedad de los grupos de ser abelianos, que se conserva por paso a un subgrupo o al producto de varios grupos, se conserva también por paso al cociente.

Tomemos de nuevo el grupo abeliano G y el subgrupo H . Entonces la aplicación canónica $p : G \rightarrow G/H$ es un homomorfismo de grupos, según resulta de (2).

El núcleo de p es evidentemente H . Por aplicación del teorema II.1.2, se obtiene el siguiente teorema, que es un caso particular del teorema II.6.2.

TEOREMA II.5.3 (*Descomposición canónica de los homomorfismo de grupos.*)

Sean G un grupo abeliano, L un subgrupo cualquiera, y $f : G \rightarrow L$ un homomorfismo de grupos. Se designa por H el núcleo de f , por $p : G \rightarrow G/H$ la aplicación canónica, por $j : f(G) \rightarrow L$ la inyección canónica y por $\bar{f} : G/H \rightarrow f(G)$ la biyección canónica. Entonces $f(G)$ es un subgrupo de L , y \bar{f} es un isomorfismo del grupo cociente G/H en el grupo $f(G)$:

$$\begin{array}{ccc} G & \xrightarrow{f} & L \\ p \downarrow & & \uparrow j \\ G/H & \xrightarrow{\bar{f}} & f(G) \end{array}$$

Ejemplos

1) Sea a un entero > 0 y $a\mathbb{Z}$ el grupo de los enteros relativos múltiplos de a . Al grupo cociente $\mathbb{Z}/a\mathbb{Z}$ se le llama *grupo de los enteros módulo a* .

Designaremos por $\chi : \mathbb{Z} \rightarrow \mathbb{Z}/a\mathbb{Z}$ al homomorfismo canónico.

Probaremos que la restricción χ_a de χ al conjunto \mathbf{N}_{a-1} es una biyección: para $A \in \mathbb{Z}/a\mathbb{Z}$ y $m \in A$, la división euclídea de m por a da:

$$m = aq + r, \quad 0 \leq r < a.$$

Puesto que $aq \in a\mathbb{Z}$, y que $a\mathbb{Z}$ es el núcleo de χ , se tiene:

$$\chi(m) = \chi(aq) + \chi(r) = \chi(r).$$

Luego χ_a es epiyectiva. Además, la relación $\chi_a(r_1) = \chi_a(r_2)$ implica

$$\chi(r_1 - r_2) = 0, \quad r_1 - r_2 \in a\mathbb{Z}, \quad \text{de donde } r_1 = r_2$$

puesto que $0 \leq r_1 \leq a-1$ y $0 \leq r_2 \leq a-1$; luego χ_a es inyectiva.

De esto resulta que $\mathbf{Z}/a\mathbf{Z}$ posee a elementos, a saber $\chi(0), \chi(1), \dots, \chi(a-1)$. Por ejemplo, la tabla del grupo $\mathbf{Z}/4\mathbf{Z}$ (en donde, para abreviar, hemos designado $\chi(u)$ por \bar{u}) es la siguiente

	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

(en la intersección de la línea \bar{i} y de la columna \bar{j} figura el elemento $\bar{i} + \bar{j}$).

Ahora vamos a determinar los *generadores* de $\mathbf{Z}/a\mathbf{Z}$, es decir, los elementos $X \in \mathbf{Z}/a\mathbf{Z}$ tales que $\text{gr}(X) = \mathbf{Z}/a\mathbf{Z}$; $\chi(1)$ es un generador de $\mathbf{Z}/a\mathbf{Z}$, luego X es un generador si, y sólo si, existe un $n \in \mathbf{Z}$ tal que $n \cdot X = \chi(1)$. Haciendo $X = \chi(x)$, la relación $nX = \chi(1)$ equivale a la congruencia entre enteros $nx \equiv 1 \pmod{a}$. Luego $\chi(x)$ engendra $\mathbf{Z}/a\mathbf{Z}$ si, y sólo si, existen enteros n y p tales que

$$nx + pa = 1.$$

Según el teorema de Bezout, esto significa que x y a son primos entre sí. Luego, los generadores de $\mathbf{Z}/a\mathbf{Z}$ son los elementos $\chi(x)$, en donde $0 \leq x \leq a-1$, y en donde x es primo con a .

2) Sea $x \in L$, en donde L es un grupo cualquiera (denotado multiplicativamente). Designemos por ψ el homomorfismo de \mathbf{Z} en L definido por $\psi(n) = x^n$; si x es de orden infinito, ψ es inyectiva, y $\text{gr}(x)$ es isomorfo a \mathbf{Z} . Si x es de orden finito $\omega > 0$, el núcleo de ψ es $\omega\mathbf{Z}$. Según el teorema II.5.3 se tiene la descomposición canónica de ψ :

$$\begin{array}{ccc} \mathbf{Z} & \xrightarrow{\psi} & L \\ \chi \downarrow & & \uparrow j \\ \mathbf{Z}/\omega\mathbf{Z} & \xrightarrow{\bar{\psi}} & \text{gr}(x) \end{array}$$

en donde $\bar{\psi}$ es un isomorfismo de grupos.

De lo que resulta que $\text{gr}(x)$ es isomorfo a $\mathbf{Z}/\omega\mathbf{Z}$.

DEFINICIÓN II.5.1

§ Sea L un grupo. Si existe un $x \in L$ tal que $\text{gr}(x) = L$, se dice que L es
 § **cíclico**, engendrado por x .

Un grupo cíclico es, pues, isomorfo a \mathbf{Z} , o bien a

$$\mathbf{Z}/\omega\mathbf{Z}, \quad \text{con } \omega \in \mathbf{N}^*.$$

3) Sea U el grupo multiplicativo de los números complejos de módulo 1. La aplicación $f_1: \theta \mapsto e^{i\theta}$ de \mathbf{R} en U es un homomorfismo de grupos *epiyectivo*, cuyo núcleo es el grupo $2\pi\mathbf{Z}$ de los múltiplos enteros de 2π . Igualmente, para todo número real $a \neq 0$, la aplicación $f_a: \theta \mapsto e^{ia\theta}$ es un homomorfismo *epiyectivo*, cuyo núcleo es $\frac{2\pi}{a}\mathbf{Z}$. Aplicando el teorema II. 5. 3 vemos que el grupo $\mathbf{R}/\frac{2\pi}{a}\mathbf{Z}$ es isomorfo

al grupo U .

Recíprocamente, se puede probar que todo homomorfismo no constante y *continuo* de \mathbf{R} en U es de la forma f_a para un número real $a \neq 0$ (cf. tomo 2).

* 4) Sea G un grupo cíclico, designado multiplicativamente, cuyo x es un generador. Si el orden de x es infinito, G es isomorfo a \mathbf{Z} , del cual conocemos los subgrupos y los grupos cocientes. Si no, sea ω el orden de x , y $\psi: \mathbf{Z} \rightarrow G$ el homomorfismo canónico, que se factoriza en

$$\mathbf{Z} \xrightarrow{\gamma} \mathbf{Z}/\omega\mathbf{Z} \xrightarrow{\bar{\psi}} G,$$

en donde $\bar{\psi}$ es un isomorfismo.

Sea, entonces, H un subgrupo de G ; la imagen recíproca $\psi^{-1}(H)$ es un subgrupo de \mathbf{Z} , por lo tanto un grupo de la forma $d\mathbf{Z}$. Puesto que $d\mathbf{Z} \supset \omega\mathbf{Z}$, d divide a ω . Pero (al ser ψ *epiyectiva*) $\psi(\psi^{-1}(H)) = H$, luego H está formado por los elementos x^{dn} , $n \in \mathbf{Z}$. Esto prueba, ante todo, que H es cíclico, engendrado por x^d . Entonces, aplicando II.5.3, se observa que H es isomorfo al grupo cociente $d\mathbf{Z}/\omega\mathbf{Z}$. Para estudiar este cociente, consideremos la aplicación $n \mapsto nd$, que es un isomorfismo de \mathbf{Z} en $d\mathbf{Z}$, y la imagen recíproca de $\omega\mathbf{Z}$ por este isomorfismo es $\frac{\omega}{d}\mathbf{Z}$. Luego $d\mathbf{Z}/\omega\mathbf{Z}$ es isomorfo a $\mathbf{Z}/\frac{\omega}{d}\mathbf{Z}$, y esto prueba que el orden de x^d es $\frac{\omega}{d}$.

Consideremos ahora con las notaciones anteriores, el grupo cociente G/H : es cíclico, engendrado por la clase de x , que designamos por X . La relación $X^n = (\text{elemento neutro de } G/H)$ equivale a: $x^n \in H$, o sea n es múltiplo de d . Dicho de otra manera, G/H es cíclico de orden d .

5) Consideremos de nuevo el grupo U del ejemplo 3), y sea $f: \mathbf{Z} \rightarrow U$ un homomorfismo; $f(\mathbf{Z})$ es un grupo cíclico engendrado por $\zeta = f(1)$. Con precisión, existe un número real $a \in [0, 2\pi]$ tal que $\zeta = e^{ia}$ (cf. ejemplo 3)), y se tiene entonces $f(n) = e^{ina}$, para todo entero n . Se presentan, pues, dos casos:

1.º Si $a \in 2\pi\mathbf{Q}$ (subgrupo de \mathbf{R} formado por los múltiplos racionales de 2π), el núcleo N de f no está reducido a $\{0\}$, y $N = d\mathbf{Z}$ para un entero $d > 0$. Luego ζ es de

orden d , la relación $\zeta^n = 1$ equivale a: n es múltiplo de d , y $f(\mathbf{Z})$ es isomorfo a $\mathbf{Z}/d\mathbf{Z}$. Se dice que ζ es una *raíz primitiva d -ésima de 1*.

2.º Si $a \notin 2\pi\mathbf{Q}$, f es inyectivo, y $f(\mathbf{Z})$ es isomorfo a \mathbf{Z} .

Recíprocamente, si ζ es un elemento de orden d de U , el subgrupo engendrado por ζ en U es finito, contiene d elementos y es isomorfo a $\mathbf{Z}/d\mathbf{Z}$. Veremos más adelante que todo subgrupo finito de U es de esta forma.

Un teorema importante que concierne a los grupos cíclicos finitos es el siguiente (generalización del ejemplo 1):

TEOREMA II.5.4

Sea G un grupo (designado multiplicativamente, y con elemento neutro e) cíclico finito, de cardinal $n \geq 2$, y engendrado por x . Entonces para todo $\alpha \in \mathbf{Z}$, el orden d de x^α es igual a

$$n' = n/\text{mcd}(\alpha, n)$$

($\text{mcd}(\alpha, n)$ designa aquí al entero > 0).

Demostración.

- a) Sea $\alpha' = \alpha/\text{mcd}(\alpha, n)$; entonces $(x^\alpha)^{n'} = x^{\alpha'n} = e$, luego n' es múltiplo de d .
 b) Hagamos $\delta = \text{mcd}(\alpha, n)$, entonces $\alpha = \alpha'\delta$, $n = n'\delta$, y α' y n' son primos entre sí.

Si el entero m es tal que $(x^\alpha)^m = e$, se deduce:

$$\alpha m \equiv 0 \pmod{n} \text{ de donde } \alpha' m \equiv 0 \pmod{n'}.$$

Puesto que α' y n' son primos entre sí, esta última relación implica: $m \equiv 0 \pmod{n'}$. c.q.d.

§ II.6 GRUPOS CUALESQUIERA: CLASES, SUBGRUPOS NORMALES. COCIENTE

Sea G un grupo (designado multiplicativamente), y H un subgrupo de G . Se pueden definir dos relaciones de equivalencia vinculadas con H , que relacionan los elementos x e y de G :

- la relación: $xy^{-1} \in H$, a cuyas clases se les llama *clases por la derecha mod(H)*;
 - la relación: $x^{-1}y \in H$, a cuyas clases se les llama *clases por la izquierda mod(H)*.
- (Dejamos al lector el trabajo de demostrar que son relaciones de equivalencia.)

1) Designamos por $(G/H)_d$ y $(G/H)_g$ al conjunto de clases por la derecha y al conjunto de clases por la izquierda. Sea A la biyección $x \mapsto x^{-1}$ de G en G . Si $xy^{-1} \in H$, se tiene $yx^{-1} \in H$, es decir, $A(y^{-1})A(x) \in H$, o

$$(A(y))^{-1} A(x) \in H.$$

Luego la imagen por A de la clase por la derecha de $x \bmod(H)$ es la clase por la izquierda de $A(x) \bmod(H)$; luego la imagen por A de toda clase por la derecha [resp. por la izquierda] es una clase por la izquierda [resp. por la derecha]. Dicho de otra manera: la aplicación $x \mapsto A(x)$ induce una biyección de $(G/H)_d$ en $(G/H)_g$.

En particular, si uno de estos conjuntos es finito, el otro también lo será, y ambos tendrán el mismo número de elementos. A este número se le llama *índice de H en G* y se designa por $[G:H]$.

Nota. $[G:\{e\}]$ es, precisamente, el cardinal $[G]$ de G siempre que G sea finito.

Para cada una de estas relaciones de equivalencia, la clase del elemento neutro e de G es H . Sin embargo, en general, ninguna de estas relaciones es compatible con la ley de G .

Analícemos, por ejemplo, bajo qué condición es compatible la primera relación.

Si $x'x^{-1} \in H$ e $y'y^{-1} \in H$, deberá verificarse $x'y'(xy)^{-1} \in H$, luego

$$(1) \quad x' y' y^{-1} x^{-1} \in H.$$

En particular, si hacemos $x' = x$ e $y \in H$, $y' = e$, vemos que *para todo $x_1 \in G$ y todo $y_1 \in H$ se debe tener: $x_1 y_1 x_1^{-1} \in H$* . Recíprocamente, si se satisface esta relación, se satisfará (1), pues:

$$\begin{aligned} x' y' y^{-1} x^{-1} &= (x' y' x'^{-1}) (x' x^{-1}) (xy^{-1} x^{-1}), \\ y \quad x' y' x'^{-1} &\in H, \quad x' x^{-1} \in H, \quad xy^{-1} x^{-1} \in H. \end{aligned}$$

La relación:

$$\text{«para todo } y \in H \text{ y todo } x \in G, xyx^{-1} \in H\text{»}$$

es equivalente a:

$$\text{«para todo } y \in H \text{ y todo } x \in G, \text{ existe un } z \in H \text{ tal que } xy = zx\text{»,}$$

por lo tanto

«para todo $x \in G$, la clase por la izquierda de x coincide con la clase por la derecha de x ».

Resumiendo, podemos enunciar:

TEOREMA II.6.1

Sea H un subgrupo de G . Para que la relación de equivalencia

$$xy^{-1} \in H \quad (\text{resp. } x^{-1}y \in H)$$

sea compatible con la ley de G , es necesario y suficiente que la clase por la izquierda de cualquier $x \in G$ coincida con su clase por la derecha, o también que, para todo $x \in G$, el automorfismo interno $y \mapsto xyx^{-1}$ deje invariante a H .

En relación con esta propiedad se establece:

DEFINICIÓN II.6.1

$\left\{ \begin{array}{l} \text{A un subgrupo } H \text{ de } G \text{ que verifica las condiciones de II.6.1 se le llama} \\ \text{subgrupo } \mathbf{distinguido}, \mathbf{invariante} \text{ o } \mathbf{normal} \text{ de } G. \end{array} \right.$

Ejemplos

1) Sean G_1 y G_2 dos grupos, de elementos neutros e_1 y e_2 , y $f: G_1 \rightarrow G_2$ un homomorfismo. El núcleo N de f es un subgrupo normal de G_1 ; en efecto, para $x \in G_1$ e $y \in N$, se tiene

$$f(xyx^{-1}) = f(x) f(y) f(x^{-1}) = f(x) e_2 f(x^{-1}) = f(x) f(x^{-1}) = f(e_1) = e_1.$$

En general, la imagen recíproca de un subgrupo normal de G_2 por f es un subgrupo normal de G_1 . Por el contrario, la imagen directa por f de un subgrupo normal de G_1 no es, en general, un subgrupo normal de G_2 .

2) Sean G_1 y G_2 dos grupos, y $G = G_1 \times G_2$ el grupo producto. Identifiquemos G_1 con $G_1 \times \{e_2\}$; sea $y \in G_1$ y $u = (a, b) \in G$. Se tiene

$$uyu^{-1} = (aya^{-1}, be_2b^{-1}) = (aya^{-1}, e_2) \in G_1.$$

Luego G_1 es un subgrupo normal en $G_1 \times G_2$.

3) Si $(H_i)_{i \in I}$ es una familia de subgrupos normales de G , el grupo

$$H = \bigcap_{i \in I} H_i$$

es normal, y el grupo engendrado por $\bigcup_{i \in I} H_i$ es normal.

Cuando H es normal en G , se puede definir una ley cociente en el conjunto de las clases (por la derecha o por la izquierda) $\text{mod}(H)$. Exactamente como en la demostración de II.5.1, se demuestra que esta ley es una ley de grupo, y que el homomorfismo canónico de G en este cociente tiene a H por núcleo. De donde

DEFINICIÓN II.6.2

Si H es un subgrupo normal de G , se llama **grupo cociente** de G por H , y se designa G/H , al conjunto de las clases $\text{mod}(H)$ provisto de la ley cociente de la de G .

Además, se deduce del teorema II.1.2 el *teorema general de homomorfismo para grupos*.

TEOREMA II.6.2

Sean G_1 y G_2 dos grupos, $f : G_1 \rightarrow G_2$ un homomorfismo, y N el núcleo de f . Sea

$$\begin{array}{ccc} G_1 & \xrightarrow{f} & G_2 \\ p \downarrow & & \uparrow j \\ G_1/N & \xrightarrow{\bar{f}} & f(G_1) \end{array}$$

la factorización canónica de \bar{f} , en donde p es la aplicación canónica, \bar{f} la biyección canónica y j la inyección canónica. Entonces p y j son homomorfismos, y \bar{f} es un isomorfismo del grupo cociente G_1/N en el subgrupo $f(G_1)$.

Ejemplos

1) Sea H un subgrupo normal de G y $p : G \rightarrow G/H$ la aplicación canónica. Sea L un grupo y $f : G/H \rightarrow L$ un homomorfismo.

Entonces $f \circ p$ es un homomorfismo cuyo núcleo contiene a H . Recíprocamente, sea $g : G \rightarrow L$ un homomorfismo de núcleo N con $H \subset N$. Toda clase de $G \text{ mod}(N)$ es reunión de clases $\text{mod}(H)$, luego el valor $g(x)$ es constante en toda clase $\text{mod}(H)$. Para $X \in G/H$, designamos por $f(X)$ el valor común de los $g(x)$ para $x \in X$. Es evidente que

$$g = f \circ p,$$

y que esta relación define a f de forma única, y además f es un homomorfismo. En resumen:

Para todo homomorfismo $g : G \rightarrow L$, cuyo núcleo contenga a H , existe un único homomorfismo $f : G/H \rightarrow L$, tal que $g = f \circ p$.

2) Sea G un grupo, y sea $\text{Aut}(G)$ el grupo de los automorfismos de G . Recordemos que para todo $x \in G$, la aplicación $\sigma_x : y \mapsto xyx^{-1} = \sigma_x(y)$ es un automorfismo, y que a estos automorfismos se les llama *automorfismos internos de G* .

La aplicación $\sigma : x \mapsto \sigma_x$ de G en $\text{Aut}(G)$ es un homomorfismo. En efecto,

$$\sigma_{x_1 x_2}(u) = x_1 x_2 u x_2^{-1} x_1^{-1} = \sigma_{x_1} \circ \sigma_{x_2}(u),$$

de donde $\sigma_{x_1 x_2} = \sigma_{x_1} \circ \sigma_{x_2}$. La imagen $\mathcal{I}(G)$ de σ es el subgrupo de los automorfismos internos. El núcleo C de σ es el conjunto de los $x \in G$ tales que para todo $y \in G$, se tiene: $xyx^{-1} = y$, o sea, $xy = yx$. Es, pues, el conjunto de los $x \in G$ que conmutan con todos los elementos de G . A este grupo C se llama **centro** de G ; puesto que C es el núcleo de σ , C es normal, e $\mathcal{I}(G)$ es isomorfo al cociente G/C .

§ II.7 GRUPOS FINITOS. GRUPO SIMÉTRICO

Si G es un grupo finito, todo subgrupo H de G es finito y de índice finito (cf. § 6, inicio). Se tiene entonces el siguiente:

TEOREMA II.7.1

Si G es un grupo finito y H un subgrupo, se tiene (designando siempre el índice de H en G por $[G : H]$ y el cardinal de G por $[G]$).

$$[G] = [G : H] \cdot [H].$$

Demostración. Consideremos, por ejemplo, el conjunto $[G/H]_a$ de las clases por la derecha. Si $X \in [G/H]_a$ y $x \in X$, la aplicación $\varphi : h \mapsto hx$, de H en G , tiene su imagen contenida en X y es inyectiva. Además, si $x' \in X$, se tiene:

$$x' x^{-1} \in H,$$

luego existe un $h \in H$ tal que $x' = h \cdot x = \varphi(h)$, luego φ es una biyección de H en X . (No depende del hecho de que G sea finito.) Se deduce que todas las clases por la derecha tienen el mismo número de elementos, igual a $[H]$. c.q.d.

COROLARIOS

1) Si H es un subgrupo de G y K es un subgrupo de H , se tiene:

$$[G : K] = [G : H] [H : K].$$

(Esta fórmula es igualmente cierta si se supone únicamente que $[G : K]$ es finito.)

2) Si G es un grupo finito, y x es un elemento de G , el **orden de x divide a $[G]$** . En particular, si hacemos $n = [G]$, tenemos: $x^n = e$.

Demostremos la relación $[G : K] = [G : H] [H : K]$ suponiendo tan sólo que K es un subgrupo de índice finito de un grupo cualquiera G , y que el subgrupo H de G verifica $K \subset H \subset G$. Sea \mathcal{G}_K (resp. \mathcal{G}_H) el conjunto de las clases por la izquierda de $G \bmod (K)$ [resp. $\bmod (H)$], y sea $\mathcal{G}_{H,K}$ el conjunto de las clases por la izquierda de $H \bmod (K)$. Se tiene ante todo: $\mathcal{G}_{H,K} \subset \mathcal{G}_K$.

Para cada $\gamma \in \mathcal{G}_H$ sea \mathcal{C}_γ el conjunto de los $c \in \mathcal{G}_K$ tales que $c \subset \gamma$; se tiene: $\mathcal{C}_H = \mathcal{G}_{H,K}$, y es claro que los $(\mathcal{C}_\gamma)_{\gamma \in \mathcal{G}_H}$ forman una *partición* de \mathcal{G}_K .

Por otro lado, el cardinal de cada conjunto \mathcal{C}_γ es igual a $[H : K] = \text{card}(\mathcal{G}_{H,K})$; (en efecto, si x es un elemento cualquiera de γ , la aplicación $h \mapsto xh$, $H \rightarrow \gamma$ es una biyección de H en γ , que define una biyección de $\mathcal{G}_{H,K}$ en \mathcal{C}_γ).

Del hecho de que los $(\mathcal{C}_\gamma)_{\gamma \in \mathcal{G}_H}$ tengan todos el mismo cardinal $[H : K]$, de que formen una partición de \mathcal{G}_K , de que $\text{card}(\mathcal{G}_K) = [G : K]$ y de que $\text{card}(\mathcal{G}_H) = [H : K]$, se deduce:

$$[G : K] = [G : H] [H : K].$$

Grupos de permutaciones de un conjunto finito con n elementos

Designamos por \mathbf{N}_n^* al conjunto $\{1, 2, \dots, n\}$ de los n primeros enteros. El grupo \mathfrak{S}_n de las biyecciones de \mathbf{N}_n^* en \mathbf{N}_n^* tiene $n!$ elementos. A veces será cómodo designar a una permutación $\sigma \in \mathfrak{S}_n$ por medio de la tabla de sus valores, o sea:

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}.$$

La permutación idéntica se designará por e_n ; el producto $\sigma \circ \tau$ de dos permutaciones se designará simplemente por $\sigma\tau$ o por $\sigma \cdot \tau$.

— Sea $m > n$. La observación siguiente nos servirá en el futuro. Hemos visto (§ 4, ejemplo 3) que la aplicación $j_{m,n}: \mathfrak{S}_n \rightarrow \mathfrak{S}_m$, tal que $j_{m,n}(\sigma)$ coincide con σ en \mathbf{N}_n^* y deja fijos los restantes elementos de \mathbf{N}_m^* , es un homomorfismo de grupos inyectivo, cuya imagen es el subgrupo de \mathfrak{S}_m formado por las permutaciones que dejan fijos $n+1, n+2, \dots, m$. En particular

$$j_{m,n}(e_n) = e_m.$$

— Para $n \geq 3$, \mathfrak{S}_n no es abeliano, puesto que contiene a \mathfrak{S}_3 y \mathfrak{S}_3 no es abeliano, según lo demuestran las relaciones siguientes (cuyas primeras definen a σ_1 y a σ_2):

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \sigma_1 \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \sigma_2 \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

— Sean $i, j \in \mathbf{N}_n^*$, $i < j$. Llamaremos *trasposición de i y j* al elemento τ_{ij}^n de \mathfrak{S}_n definido por

$$\tau_{ij}^n(i) = j, \quad \tau_{ij}^n(j) = i, \quad \text{y} \quad \tau_{ij}^n(k) = k \quad \text{para} \quad k \neq i, k \neq j.$$

Una trasposición es de orden 2. Es claro que, para $m > n$, $j_{m,n}(\tau_{ij}^n) = \tau_{ji}^m$. (Para la cuestión de notaciones ver más arriba.) Se tiene entonces:

TEOREMA II.7.2

|| Si $n \geq 2$, el conjunto de las trasposiciones de \mathfrak{S}_n es un sistema de generadores de \mathfrak{S}_n .

Demostración. Para $n = 2$, $\mathfrak{S}_2 = \{e, \tau_{12}^2\}$. Supongamos cierto el teorema para el entero $n - 1$, y sea $\sigma \in \mathfrak{S}_n$.

Primer caso: $\sigma(n) = n$. En este caso, la restricción σ' de σ a \mathbf{N}_{n-1}^* es un elemento de \mathfrak{S}_{n-1} . Por la hipótesis de recurrencia, existen trasposiciones τ'_1, \dots, τ'_p de \mathfrak{S}_{n-1} tales que $\sigma' = \tau'_1 \tau'_2 \dots \tau'_p$. Sea $\tau_i = j_{n-1,n}(\tau'_i)$; entonces τ_i es una trasposición de \mathfrak{S}_n y se tiene $\sigma = \tau_1 \tau_2 \dots \tau_p$.

Segundo caso: $\sigma(n) = q < n$. Hagamos: $\tau = \tau_{q,n}^n$, y $\psi = \tau \cdot \sigma$.

Se tiene: $\psi(n) = \tau(\sigma(n)) = \tau(q) = n$. Según el primer caso existen trasposiciones $\tau_1 \dots \tau_p$ tales que $\psi = \tau_1 \dots \tau_p$, de donde

$$\tau \cdot \sigma = \tau_1 \dots \tau_p, \quad \sigma = \tau \cdot \tau_1 \dots \tau_p. \quad \text{c.q.d.}$$

La descomposición de $\sigma \in \mathfrak{S}_n$ en producto de trasposiciones no es única. Sin embargo, la paridad del número de estas trasposiciones, para σ fijo, está bien determinada; es lo que vamos a probar y precisar, introduciendo la noción de *signatura*

DEFINICIÓN II.7.1

~ Para toda permutación $\sigma \in \mathfrak{S}_n$, se llama **signatura** de σ , y se designa por $\varepsilon(\sigma)$, al número

$$\varepsilon(\sigma) = \frac{\prod_{1 \leq i < j \leq n} (\sigma(i) - \sigma(j))}{\prod_{1 \leq i < j \leq n} (i - j)}.$$

Designamos por T_n al conjunto $\{(i, j) \mid (i, j) \in \mathbf{N}_n^* \times \mathbf{N}_n^* \text{ e } i < j\}$, por $\Delta = \{(i, i)\}_{1 \leq i \leq n}$ a la *diagonal* de $\mathbf{N}_n^* \times \mathbf{N}_n^*$, y por K_n al conjunto

$$(\mathbf{N}_n^* \times \mathbf{N}_n^*) \setminus \Delta.$$

Para toda permutación $\rho \in \mathfrak{S}_n$, definimos las aplicaciones

$$f_\rho : T_n \rightarrow K_n \quad \text{y} \quad g_\rho : T_n \rightarrow T_n$$

por medio de las fórmulas:

$$f_\rho((i, j)) = (\rho(i), \rho(j)),$$

$$g_\rho((i, j)) = \begin{cases} f_\rho((i, j)) & \text{si } \rho(i) < \rho(j), \\ (\rho(j), \rho(i)) & \text{si } \rho(i) > \rho(j). \end{cases}$$

f_ρ es una inyección y g_ρ es una biyección. Reordenando el conjunto T_n por medio de g_ρ , se tiene:

$$\varepsilon(\rho) = \frac{\prod_{g_\rho((i,j)) \in T_n} (\rho(i) - \rho(j))}{\prod_{(i,j) \in T_n} (i - j)} = (-1)^N \frac{\prod_{(i,j) \in T_n} (i - j)}{\prod_{(i,j) \in T_n} (i - j)} = (-1)^N,$$

en donde N es el conjunto de los pares (i, j) tales que $\rho(i) > \rho(j)$ e $i < j$ («número de inversiones» de ρ).

Si designamos por Γ al grupo multiplicativo de los números $\{-1, +1\}$, acabamos de ver que ε envía \mathfrak{S}_n a Γ . Con más precisión:

TEOREMA II.7.3

|| La *signatura* ε es un homomorfismo del grupo \mathfrak{S}_n en el grupo multiplicativo $\Gamma = \{-1, +1\}$.

Demostración ()*. Conservamos las notaciones anteriores. Sean $\sigma, \rho \in \mathfrak{S}_n$. Por definición se tiene:

$$\begin{aligned} \varepsilon(\sigma\rho) &= \prod_{1 \leq i < j \leq n} \frac{\sigma\rho(i) - \sigma\rho(j)}{i - j} = \prod_{i < j} \frac{\sigma\rho(i) - \sigma\rho(j)}{\rho(i) - \rho(j)} \prod_{i < j} \frac{\rho(i) - \rho(j)}{i - j} \\ &= \varepsilon(\rho) A_\sigma, \quad \text{con : } A_\sigma = \prod_{i < j} \frac{\sigma\rho(i) - \sigma\rho(j)}{\rho(i) - \rho(j)}. \end{aligned}$$

(*) Si se analiza la demostración a fondo, el lector verá que consiste en establecer la propiedad siguiente: la *signatura* $\varepsilon(\sigma)$ no depende del orden total elegido en el conjunto finito \mathbf{N}_n^* para calcularla. $\varepsilon(\sigma)$ depende sólo de σ y de $n = \text{card}(\mathbf{N}_n^*)$. Para una definición correcta del cardinal de un conjunto (cf. [4]).

Para $i, j \in \mathbf{N}_n^*$, $i \neq j$, escribimos $S((i, j)) = \frac{\sigma(i) - \sigma(j)}{i - j}$, y es:

$$(1) \quad S((i, j)) = S((j, i)).$$

Se puede escribir:

$$A_\sigma = \prod_{i < j} S(f_\sigma((i, j))) ;$$

luego (en virtud de (1)), $A_\sigma = \prod_{i < j} S(g_\sigma(i, j))$, y dado que g_σ es una biyección de T_n en T_n ,

$$A_\sigma = \prod_{i < j} S((i, j)) = \prod_{i < j} \frac{\sigma(i) - \sigma(j)}{i - j} = \varepsilon(\sigma).$$

Resulta, pues: $\varepsilon(\sigma\rho) = \varepsilon(\sigma)\varepsilon(\rho)$. c.q.d.

Si $n \geq 2$, ε es epiyectiva, ya que $\varepsilon(\tau_{12}^2) = -1$. El núcleo de ε (conjunto de los $\sigma \in \mathfrak{S}_n$ tales que $\varepsilon(\sigma) = 1$) es un subgrupo de \mathfrak{S}_n , que se designa por A_n .

DEFINICIÓN II.7.2

$\left\{ \begin{array}{l} \text{Al grupo } A_n \text{ formado por las permutaciones } \sigma \in \mathfrak{S}_n \text{ tales que } \varepsilon(\sigma) = +1 \\ \text{se le llama } \mathbf{grupo\ alternado} \text{ de } n \text{ elementos. A los elementos de } A_n \\ \text{se les llama permutaciones } \mathbf{pares}, \text{ a los elementos de } \mathfrak{S}_n \setminus A_n \text{ se les} \\ \text{llama permutaciones } \mathbf{impares}. \end{array} \right.$

A_n es normal en \mathfrak{S}_n , y el grupo cociente \mathfrak{S}_n/A_n es isomorfo a $\Gamma = \{-1, +1\}$; se tiene pues

$$[\mathfrak{S}_n : A_n] = 2.$$

Probaremos ahora que *toda trasposición es una permutación impar*.

Sean $i < j$ y τ la trasposición de \mathfrak{S}_n definida sobre $\{i, j\}$. Los pares (k, l) tales que $\tau(k) > \tau(l)$ y $k < l$ son:

- los pares (i, k) con $k \leq j$, en número de $j - i$;
- los pares (k, j) con $i < k < j$, en número de $j - i - 1$.

En total tendremos $2(j - i) - 1$ pares, y este número es impar, de donde resulta nuestra afirmación.

Consecuencia. Descomponemos $\sigma \in \mathfrak{S}_n$ en un producto de trasposiciones:

$$\sigma = \tau_1 \tau_2 \dots \tau_p.$$

Se tiene:

$$\varepsilon(\sigma) = \varepsilon(\tau_1) \varepsilon(\tau_2) \dots \varepsilon(\tau_p) = (-1)^p,$$

luego la clase de $p \pmod{2}$ está bien determinada: p es par o impar según que σ sea par o impar.

§ II.8 GRUPO QUE OPERA SOBRE UN CONJUNTO

En este §, todos los grupos que intervengan se designarán multiplicativamente, y designaremos por e a su elemento neutro (salvo mención expresa).

DEFINICIÓN II.8.1

Se dice que el grupo G opera por la izquierda sobre el conjunto E , si se ha dado una ley externa por la izquierda sobre E , de dominio G , a saber

$$(g, x) \mapsto g \cdot x \quad (g \in G, x \in E),$$

que verifica las condiciones:

- (G₁) para todo $g_1, g_2 \in G$ y todo $x \in E$, $(g_1 g_2) \cdot x = g_1(g_2 \cdot x)$
- (G₂) para todo $x \in E$, $e \cdot x = x$.

Se podría definir análogamente un grupo G que opera por la derecha sobre el conjunto E , dando una ley externa por la derecha $(x, g) \mapsto x \cdot g$ en E , tal que

$$x(g_1 \cdot g_2) = (xg_1) \cdot g_2 \quad \text{y} \quad x \cdot e = x \quad \text{para } x \in E, \quad g_1, g_2 \in G.$$

En lo que sigue nos ocuparemos esencialmente de los grupos que operan por la izquierda sobre un conjunto.

Designemos por \mathfrak{S}_E al grupo de las biyecciones de E en E ; fijemos $g \in G$, y sea τ_g la aplicación $x \mapsto g \cdot x$ de E en E . En virtud de (G₁) y (G₂) se tiene:

$$(\tau_{g^{-1}} \circ \tau_g)(x) = g^{-1}(g \cdot x) = (g^{-1} \cdot g) \cdot x = e \cdot x = x = (\tau_g \circ \tau_{g^{-1}})(x).$$

Luego τ_g y $\tau_{g^{-1}}$ son biyecciones inversas la una de la otra.

Además, si $g_1, g_2 \in G$,

$$(\tau_{g_1} \circ \tau_{g_2})(x) = g_1(g_2 \cdot x) = (g_1 g_2) \cdot x = \tau_{g_1 g_2}(x).$$

En otras palabras, la aplicación $g \mapsto \tau_g$ de G en \mathfrak{S}_E es un homomorfismo de grupos. Recíprocamente, si $g \mapsto \tau_g$ es un homomorfismo de G en \mathfrak{S}_E , la ley externa por la izquierda en E definida por $(g, x) \mapsto \tau_g(x) = g \cdot x$ hace operar a G por la izquierda sobre E . De donde:

II.8.1 Dar una operación por la izquierda $(g, x) \mapsto g \cdot x$ del grupo G sobre el conjunto E equivale a dar un homomorfismo $g \mapsto \tau_g$ de G en \mathfrak{S}_E , definido por $\tau_g(x) = g \cdot x$ para $g \in G$ y $x \in E$.

Subgrupo de isotropía. Órbitas

Sea G un grupo que opera por la izquierda sobre E , y sea $a \in E$. El conjunto G_a de los $g \in G$ tales que $g.a = a$ es un subgrupo de G , pues, si $g_1, g_2 \in G_a$, se tiene:

$$g_1 g_2 \in G_a$$

según (G_1) ; y si $g \in G_a$, se tiene

$$g^{-1}.a = g^{-1}.(g.a) = (g^{-1}g).a = e.a = a$$

según (G_2) , de donde $g^{-1} \in G_a$; finalmente, G_a es no vacío, ya que (según (G_2)) $e \in G_a$.

DEFINICIÓN II.8.2

Si el grupo G opera por la izquierda sobre el conjunto E , al subgrupo G_a de los $g \in G$ que dejan fijo un elemento $a \in E$ se le llama **subgrupo de isotropía** de a .

La relación binaria definida en E por:

«existe un $g \in G$ tal que $y = g.x$ » es una *relación de equivalencia*.

En efecto, es reflexiva, según (G_2) . Según (G_1) , $y = g.x$ implica

$$g^{-1}.y = g^{-1}.(g.x) = e.x = x,$$

luego la relación es simétrica. Finalmente, si $y = g.x$ y $z = h.y$ ($g \in G$, $h \in G$), se tiene:

$$z = h.(g.x) = (hg).x,$$

luego la relación es transitiva. Podemos enunciar:

II.8.2 Si el grupo G opera por la izquierda sobre el conjunto E , la relación «existe $g \in G$ tal que $y = g.x$ ($x, y \in E$)» es una *relación de equivalencia* sobre E , cuyas clases se llaman **órbitas de E según G** , o **G -órbitas de E** .

DEFINICIÓN II.8.3

Se dice que G **opera transitivamente** sobre E si el número de órbitas según G es igual a 1; en otras palabras si, para todo $x \in E$ y todo $y \in E$, existe un $g \in G$ tal que $y = g.x$. En caso contrario, se dice que G **opera intransitivamente** sobre E .

Vamos a interesarnos ahora por los subgrupos de isotropía de los elementos de una órbita fija. Precisemos que dos subgrupos H_1 y H_2 de un grupo G son *conjugados* si se transforman el uno en el otro por medio de un *automorfismo interno* de G ; en otros términos, si existe un $\sigma \in G$ tal que $H_2 = \sigma^{-1}H_1\sigma$.

TEOREMA II.8.3

|| Sea Ω una órbita según G , el grupo G opera por la izquierda sobre el conjunto E . Si $a \in \Omega$ y $b \in \Omega$, los subgrupos de isotropía G_a y G_b son conjugados en G .

Demostración. Sea $\sigma \in G$ tal que $b = \sigma \cdot a$. La relación $g \cdot b = b$ ($g \in G$) equivale a: $(g \cdot \sigma)a = \sigma \cdot a$, o aún a $(\sigma^{-1}g\sigma) \cdot a = a$.

Esto demuestra que la aplicación $g \mapsto \sigma^{-1}g\sigma$ es una biyección de G_b en G_a , es decir que: $G_a = \sigma^{-1} \cdot G_b \cdot \sigma$. ||

Nota. Decir que G_a es un subgrupo normal de G significa que $G_a = G_b$ para todo $a \in \Omega$ y todo $b \in \Omega$, o también, que toda aplicación τ_g que deja fijo un punto a de Ω se reduce a la identidad en Ω .

TEOREMA II.8.4

|| El grupo G opera por la izquierda sobre el conjunto E , Ω es una órbita según G , y $a \in \Omega$. Para todo $x \in \Omega$, designamos por C_x al conjunto de los $\sigma \in G$ tales que $\sigma \cdot a = x$; la aplicación $x \mapsto C_x$ es una biyección de Ω en el conjunto de clases por la izquierda de G según el subgrupo de isotropía G_a .

Demostración. Para todo $x \in \Omega$, sea $\sigma_0 \in G$ un elemento particular tal que

$$\sigma_0 \cdot a = x.$$

La relación $\sigma \cdot a = x$ (en donde $\sigma \in G$) se escribe $\sigma \cdot a = \sigma_0 \cdot a$ y equivale a $(\sigma_0^{-1} \cdot \sigma)a = a$, es decir: $\sigma_0^{-1} \cdot \sigma \in G_a$. Las relaciones:

$$\ll \sigma \in C_x \gg \text{ y } \ll \sigma \in \sigma_0 \cdot G_a \gg$$

son pues equivalentes, lo que prueba que $C_x = \sigma_0 \cdot G_a$. c.q.d.

Aplicación

Cuando la órbita Ω es un conjunto *finito*, el número entero $\text{card}(\Omega)$ es igual al índice de G_a en G , o sea a $[G : G_a]$. De ahí la ecuación

$$(1) \quad \text{card}(\Omega) = [G : G_a].$$

Hemos encontrado de nuevo el hecho de que $[G : G_a]$ es independiente de $a \in \Omega$, que resulta también de II.8.3.

Si el conjunto E es, a su vez, finito, cada órbita es un conjunto finito. Designemos por \mathcal{C} una parte de E que contenga *un elemento y sólo uno de cada órbita*. La ecuación (1), aplicada a cada órbita, nos da la relación importante, llamada *ecuación de clases*:

$$(2) \quad \text{card}(E) = \sum_{a \in \mathcal{C}} [G : G_a].$$

(En efecto, $\text{card}(E)$ es la suma de los $\text{card}(\Omega)$ para las diversas órbitas según G , ya que estas órbitas definen una partición de E .)

Grupos que operan fielmente

DEFINICIÓN II.8.4

Sea G un grupo que opera por la izquierda sobre un conjunto E . Se dice que G opera fielmente sobre E si las relaciones:

$$\sigma \in G, \text{ y } \sigma.x = x \text{ para todo } x \in E, \text{ implican: } \sigma = e.$$

Equivale a decir que el homomorfismo $g \mapsto \tau_g$ de G en \mathfrak{S}_E asociado canónicamente a la operación de G en E es inyectivo. Cuando el grupo G opera fielmente sobre E , se puede identificar G con un subgrupo de \mathfrak{S}_E .

Supongamos que G opera fiel y transitivamente sobre E ; según la nota que sigue a II.8.3, si E tiene más de un elemento y si existe un subgrupo de isotropía $G_a \neq \{e\}$, ($a \in E$), G_a no es un subgrupo normal de G ; en particular si G es abeliano, todo subgrupo de isotropía se reduce a $\{e\}$.

Vamos ahora a aplicar a algunos ejemplos simples las nociones abstractas dadas anteriormente.

Ejemplo 1. Extensión a las partes de E . Clases de p -transitividad

Si el grupo G opera por la izquierda sobre el conjunto E , para toda parte A de E , y todo $g \in G$, designemos por $g.A = \tau_g(A)$ el conjunto de los $g.a$ para $a \in A$. Se verifican trivialmente las condiciones (G_1) y (G_2) , de forma que la ley externa $A \mapsto g.A$, de dominio G , define una operación por la izquierda de G en $\mathcal{P}(E)$, llamada *extensión* a $\mathcal{P}(E)$ (de la operación por la izquierda de G en E). Si \mathcal{F} designa una parte de $\mathcal{P}(E)$, estable por la ley externa $A \mapsto g.A$ (dicho de otra manera: si \mathcal{F} es una reunión de órbitas de $\mathcal{P}(E)$ según G), la restricción a \mathcal{F} de la ley externa define una operación por la izquierda de G en \mathcal{F} . Por ejemplo, se puede tomar $\mathcal{F} = \mathcal{P}_p(E)$, conjunto de las partes con p elementos de E ($p \in \mathbb{N}^*$).

Se puede considerar también el conjunto $\mathcal{D}_p(E)$ de las aplicaciones *inyectivas* de \mathbb{N}_p^* en E : se hace operar a G sobre \mathcal{D}_p «por extensión» asociando, a todo elemento $g \in G$ y todo (x_1, x_2, \dots, x_p) , de $\mathcal{D}_p(E)$, el elemento $(y_1, \dots, y_p) \in \mathcal{D}_p(E)$ tal que $y_i = g.x_i$ para $i = 1, 2, \dots, p$.

Al conjunto de las órbitas de $\mathcal{D}_p(E)$ según G se le llama *conjunto de las clases de p -transitividad* según G ; y se dice que G opera p veces transitivamente sobre E , si G opera transitivamente sobre $\mathcal{D}_p(E)$.

Volviendo a la extensión global (de la operación de G sobre E) a $\mathcal{P}(E)$, observamos que, para toda parte no vacía A de E , el grupo de isotropía G_A es igual al conjunto de los $g \in G$ tales que τ_g deja a A globalmente invariante.

Ejemplo 2

Sea H un subgrupo cualquiera del grupo G . Hacemos que H opere por la izquierda sobre G por traslaciones por la izquierda, definiendo sobre G la ley externa de dominio H , por:

$$(h, x) \mapsto h.x \quad (h \in H, x \in G).$$

Es evidente que las órbitas de G son exactamente las *clases por la derecha* según H . Cuando G es finito la fórmula (2) se reduce a:

$$[G] = [G : H] [H].$$

(*) Aquí el autor utiliza « G opera propiamente» para indicar que « G opera fielmente» (def. II.8.4)

Ejemplo 3. Clases de conjugación. Normalizador

Hagamos que el grupo G opere sobre sí mismo por la izquierda por automorfismos internos, definiendo

$$\sigma \cdot x = \sigma x \sigma^{-1} \quad \text{para } x \in G, \sigma \in G.$$

A las órbitas de G según G se les llaman *clases de conjugación de los elementos de G* .

Un automorfismo interno de G transforma un subgrupo en un subgrupo. Podemos por tanto considerar la extensión de esta operación al conjunto \mathcal{G} de los subgrupos de G . (Brevemente, se dice que se hace operar a G sobre \mathcal{G} por automorfismos internos.) Al grupo de isotropía de un subgrupo H de G (considerado como elemento de \mathcal{G}) por la acción de G sobre \mathcal{G} , se le llama *normalizador* de H . Lo designaremos por N_H : es el conjunto de los elementos $\sigma \in G$ tales que $\sigma H \sigma^{-1} = H$.

Manifiestamente $H \subset N_H$ y H es *normal* en N_H . Además, todo subgrupo G' de G en el que $H \subset G'$ con H normal en G' es tal que $G' \subset N_H$.

Cuando G es finito (en cuyo caso \mathcal{G} también es finito), la ecuación (1) da una expresión del número ν_H de los subgrupos conjugados del subgrupo H :

$$\nu_H = [G : N_H].$$

Ejemplo 4. Descomposición de una permutación en ciclos

Designemos por \mathfrak{S}_n al grupo simétrico de orden n , formado por las biyecciones de \mathbf{N}_n^* en sí mismo. Todo subgrupo G de \mathfrak{S}_n opera por la izquierda sobre \mathbf{N}_n^* , por la fórmula

$$s \cdot x = s(x) \quad (s \in G, x \in \mathbf{N}_n^*).$$

Vamos a hacer operar sobre \mathbf{N}_n^* el subgrupo cíclico G de \mathfrak{S}_n , engendrado por una permutación $s \in \mathfrak{S}_n$, $s \neq e$. Una órbita Ω de \mathbf{N}_n^* según G se reduce a un elemento a si, y sólo si, se tiene: $s(a) = a$. Sea Ω una órbita no reducida a un elemento, y hagamos

$$p = \text{card}(\Omega) \quad (2 \leq p \leq n).$$

El orden de s es igual a $[G]$, y lo designaremos por ω ; según (1), p divide a ω : hacemos $\omega = pq$. Designamos por a a un elemento fijo de Ω .

El subgrupo de isotropía G_a es un subgrupo de G de orden q . Luego, al ser G cíclico de orden ω , posee *un solo* subgrupo de orden q , a saber, el grupo engendrado por s^p (ver def. II.5.1). Dicho de otra manera, se tiene:

$$s^p(a) = a, \quad \text{y} \quad s^k(a) \neq a \quad \text{para } k \leq p-1.$$

Al ser válido para todo $a \in \Omega$, resulta que los elementos

$$a = s^0(a), s(a), \dots, s^{p-1}(a)$$

son distintos dos a dos; puesto que hay p , son *los* elementos de Ω . Si hacemos $a_1 = a$, $a_2 = s(a)$, \dots , $a_p = s^{p-1}(a)$, la restricción de s a Ω se puede escribir en la forma:

$$\begin{pmatrix} a_1 & a_2 & \dots & a_{p-1} & a_p \\ a_2 & a_3 & \dots & a_p & a_1 \end{pmatrix}.$$

Se da en consecuencia la siguiente definición:

DEFINICIÓN II.8.5

Sea $s \in \mathfrak{S}_n$ (en donde el entero n es ≥ 2) y sea G_s el subgrupo cíclico de \mathfrak{S}_n engendrado por s . Se dice que s es un **ciclo** si $s \neq e$, y si existe, para la operación natural de G_s sobre \mathbf{N}_n^* , una órbita Ω y sólo una, no reducida a un solo elemento. Al número $l = \text{card}(\Omega)$ se le llama **longitud** del ciclo s .

A la parte Ω de \mathbf{N}_n^* se le llama *soporte* de s .

El estudio que precede se puede resumir y completar en el teorema siguiente:

TEOREMA II.8.5

Con las notaciones de la definición II.8.5, el **orden** del ciclo $s \in \mathfrak{S}_n$ es igual a su **longitud**.

Demostración. Sea $a \in \Omega$; si existen dos enteros m_1, m_2 tales que $m_2 - m_1 < l$, $0 \leq m_1 \leq m_2 \leq l$ y $s^{m_1}(a) = s^{m_2}(a)$, se deducirá (haciendo $m = m_2 - m_1$): $s^m(a) = a$. El conjunto $E = \{a, s(a), \dots, s^{m-1}(a)\}$, contenido en Ω , será G_s -estable, por tanto reunión de órbitas según G_s , de donde resulta una contradicción ya que $\text{card}(E) \leq l - 1$. Luego, $a, s(a), \dots, s^{l-1}(a)$ son distintos, luego: $\Omega = \{a, s(a), \dots, s^{l-1}(a)\}$ y necesariamente $s^l(a) = a$ (puesto que $s^l(a) \in \Omega$). De todo ello resulta que $\forall a \in \Omega$ $s^l(a) = a$, de donde se deduce que $l \equiv 0 \pmod{\omega}$, siendo ω el orden de s ; si se tuviera $\omega < l$, sería $s^\omega(a) = a$, en contradicción con la propiedad anterior. Luego $\omega = l$. c.q.d.

Dos ciclos son *disjuntos* si sus soportes lo son; dos de tales ciclos, evidentemente, *conmutan*.

Si consideramos la operación natural de G_s sobre \mathbf{N}_n^* , introducida anteriormente, la teoría de grupos que operan sobre un conjunto permite establecer:

TEOREMA II.8.6

Sea $s \in \mathfrak{S}_n$, $s \neq e$ (en donde $n \in \mathbf{N}_n^*$ y $n \geq 2$). Sea Γ_s el conjunto de las G_s -órbitas de \mathbf{N}_n^* para la operación natural de G_s , no reducidas a un solo elemento. Entonces existe una familia única $(c_\gamma)_{\gamma \in \Gamma_s}$ de ciclos, en que, para todo $\gamma \in \Gamma_s$, c_γ es un ciclo de soporte γ , y tal que

$$(3) \quad s = \prod_{\gamma \in \Gamma_s} c_\gamma$$

(los c_γ conmutan dos a dos). Además, el orden de s es igual al mcm de los órdenes de los c_γ .

Demostración (esbozo)

Para todo $\gamma \in \Gamma_s$, c_γ está necesariamente definido por la *restricción* de s a γ : basta con asegurarse de que cada uno de los c_γ así definidos es un ciclo, lo que resulta fácilmente del hecho de que γ no contiene ningún subconjunto no vacío, G_s -estable, distinto de sí mismo. Puesto que los ciclos c_γ son disjuntos, conmutan dos a dos, de donde se sigue (3).

Queda por ver que, si ω es el orden de s , y, para todo $\gamma \in \Gamma_s$, ω_γ es el orden de c_γ , se tiene: $\omega = \text{mcm}(\omega_\gamma)_{\gamma \in \Gamma_s}$.

El mcm (ω_γ) es evidentemente múltiplo de ω . Recíprocamente, la relación $s^\omega = e$ implica evidentemente, por restricción, $(c_\gamma)^\omega = e$, para todo $\gamma \in \Gamma_s$, de donde

$$\forall \gamma \in \Gamma_s \quad \omega \equiv 0 \pmod{(\omega_\gamma)}.$$

Luego ω es múltiplo del mcm (ω_γ) . c.q.d.

Ejemplo 5. Grupo de las isometrías que dejan invariante una figura.

Empecemos por estudiar el grupo \mathcal{G} de las isometrías planas que dejan globalmente invariantes los tres vértices A, B, C de un triángulo equilátero.

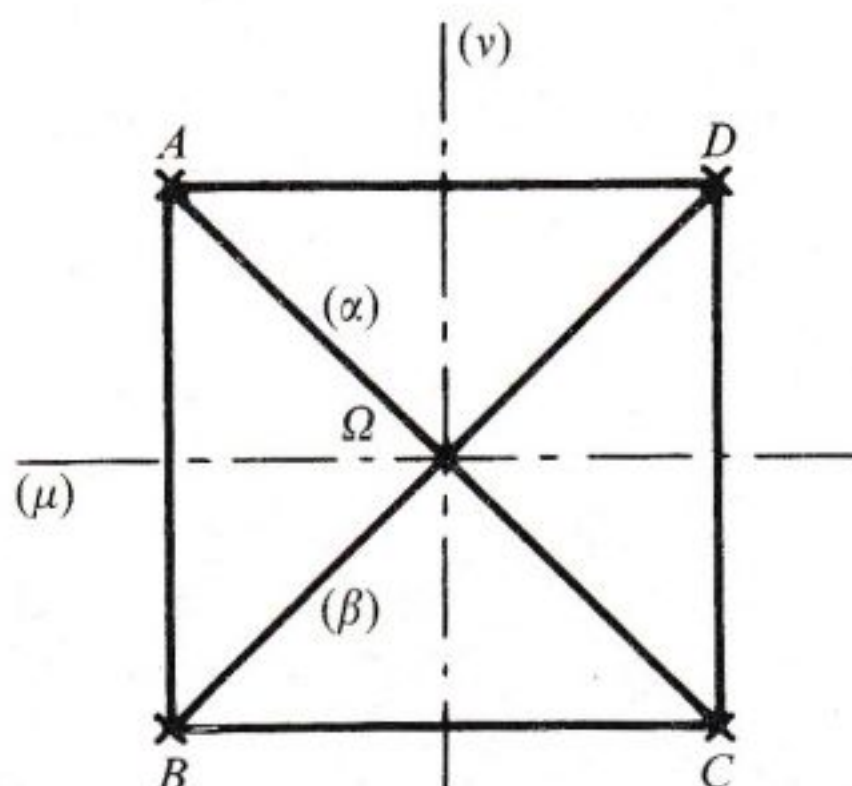
\mathcal{G} opera por la izquierda sobre el conjunto $T = \{A, B, C\}$, por la fórmula

$$f \cdot A = f(A) \quad (f \in \mathcal{G}).$$

\mathcal{G} opera fielmente, ya que una isometría plana que deja fijos tres puntos no alineados es la identidad (cf. Cap. XIII). Por lo tanto \mathcal{G} se identifica con un subgrupo de \mathfrak{S}_3 , grupo de las permutaciones de T . Pero \mathcal{G} contiene las simetrías respecto de las tres alturas de T , y estas simetrías corresponden a las *trasposiciones* sobre T . Puesto que el conjunto de las trasposiciones genera \mathfrak{S}_3 , vemos que $\mathcal{G} = \mathfrak{S}_3$. Sabemos que las permutaciones pares son los productos de un número par de trasposiciones, y que las isometrías pares son los productos de un número par de simetrías. Por consiguiente, el grupo de los *desplazamientos* que dejan invariante a T es isomorfo al grupo de las permutaciones pares sobre T , es decir, al grupo alternado \mathcal{A}_3 .

De forma general, se puede definir el *símplex regular* T_n de \mathbf{R}^n : es la figura formada por $n + 1$ puntos de la esfera unidad, equidistantes dos a dos. La existencia de T_n se puede demostrar por recurrencia sobre n (cf. ejercicios): así, T_3 es el *tetraedro regular*. Por un razonamiento completamente análogo al precedente, se puede ver que el grupo de *isometrías* de \mathbf{R}^n que dejan globalmente invariante a T_n se identifica con el grupo simétrico \mathfrak{S}_{n+1} , y el grupo de *desplazamientos* de \mathbf{R}^n que dejan invariante a T_n se identifica con el grupo alternado \mathcal{A}_{n+1} (cf. ejercicios, y también [1]).

Estudiemos ahora el grupo de isometrías planas que dejan globalmente invariantes los cuatro vértices (A, B, C, D) de un cuadrado de centro Ω . Como antes, podemos hacer que \mathcal{G} opere sobre el conjunto $Q = \{A, B, C, D\}$, y vemos que \mathcal{G} opera fielmente. Luego, \mathcal{G} se identifica con un subgrupo de \mathfrak{S}_4 .



Hagamos que \mathcal{D} opere sobre el conjunto

$$\Delta = \{ \alpha, \beta \}$$

de las dos diagonales ($\alpha = AC$, $\beta = BD$).

\mathcal{D} opera transitivamente sobre Δ , puesto que, por ejemplo, la rotación de centro Ω y ángulo $\pi/2$ cambia α y β . Las isometrías que dejan invariante cada diagonal forman un grupo \mathcal{G} de 4 elementos,

$$\mathcal{G} = \{ e, s_{\Omega}, s_{\alpha}, s_{\beta} \},$$

en donde s_{Ω} , s_{α} y s_{β} designan respectivamente las simetrías respecto a Ω , α , y β , y cuya tabla es la siguiente:

	e	s_{Ω}	s_{α}	s_{β}
e	e	s_{Ω}	s_{α}	s_{β}
s_{Ω}	s_{Ω}	e	s_{β}	s_{α}
s_{α}	s_{α}	s_{β}	e	s_{Ω}
s_{β}	s_{β}	s_{α}	s_{Ω}	e

Se observa que \mathcal{G} es conmutativo y que todos sus elementos son de orden 2; \mathcal{G} es isomorfo a $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Puesto que \mathcal{D} opera transitivamente sobre Δ , se ve que \mathcal{D} contiene exactamente 8 elementos. Los elementos de $\mathcal{G} \setminus \mathcal{D}$ son las rotaciones $(\Omega, \pm \pi/2)$ y las simetrías s_{μ} y s_{ν} respecto de las medianas μ y ν del cuadrado.

\mathcal{D} no es conmutativo: por ejemplo,

$$s_{\mu} \circ s_{\alpha} = \text{rot} \left(\Omega, + \frac{\pi}{2} \right) \quad \text{y} \quad s_{\alpha} \circ s_{\mu} = \text{rot} \left(\Omega, - \frac{\pi}{2} \right)$$

\mathcal{G} es normal en \mathcal{D} , de índice 2.

El grupo \mathcal{D} de los desplazamientos de \mathcal{D} es el grupo cíclico de 4 elementos, engendrado por $\text{rot}(\Omega, \pi/2)$, igual a $\{e, \text{rot}(\Omega, + \pi/2), s_{\Omega}, \text{rot}(\Omega, - \pi/2)\}$. Evidentemente

$$\mathcal{D} \cap \mathcal{G} = \{ e, s_{\Omega} \}.$$

Aquí se observa una diferencia esencial con el ejemplo del triángulo equilátero: \mathcal{D} no está contenido en el grupo alternado \mathcal{A}_4 . En efecto, $\text{rot}(\Omega, + \pi/2)$ engendra \mathcal{D} y corresponde a la permutación circular $\begin{pmatrix} A & B & C & D \\ B & C & D & A \end{pmatrix}$, que es *impar* (en general, una permutación circular de orden n es *impar*, y viceversa). \mathcal{G} está formado por los 4 elementos $\{e, s_{\Omega}, s_{\mu}, s_{\nu}\}$ y contiene, por lo tanto, los dos desplazamientos e, s_{Ω} , y las dos isometrías impares s_{μ}, s_{ν} . Además: $\mathcal{D} \cap \mathcal{A}_4 = \{e, s_{\Omega}\}$.

Es interesante observar que *el grupo \mathcal{G} es normal en \mathcal{A}_4* (y de índice 3). En efecto, \mathcal{G} está formado por e , y por tres productos de dos trasposiciones disjuntas de \mathfrak{S}_4 . Los otros elementos de \mathcal{A}_4 son los 8 ciclos de orden 3 de \mathfrak{S}_4 . La verificación de nuestra afirmación se hace entonces sin dificultad: en virtud de la simetría de la cuestión, basta con verificar que el conjugado $\sigma s \sigma^{-1}$ de un elemento $s \in \mathcal{G} \cap \mathcal{A}_4$ por un elemento $\sigma \in \mathcal{A}_4 \setminus (\mathcal{G} \cap \mathcal{A}_4)$ está en $\mathcal{G} \cap \mathcal{A}_4$, lo cual es fácil.

Para un mayor número de detalles acerca de los grupos finitos de isometrías de \mathbb{R}^2 y \mathbb{R}^3 , cf. [1].

Capítulo III

Estructuras algebraicas en las que intervienen varias leyes

§ III.1 GENERALIDADES

DEFINICIÓN III.1.1

Sea E un conjunto provisto de dos leyes de composición interna, designadas por \top y \perp .

La ley \top es distributiva por la izquierda respecto de la ley \perp si para todo $x, y, z \in E$, $x \top (y \perp z) = (x \top y) \perp (x \top z)$.

La ley \top es distributiva por la derecha respecto de la ley \perp si para todo $x, y, z \in E$, $(x \perp y) \top z = (x \top z) \perp (y \top z)$.

La ley \top es distributiva respecto de la ley \perp si es distributiva por la derecha y por la izquierda.

Es evidente que si \top es conmutativa, su distributividad respecto de \perp equivale a su distributividad por la derecha, o por la izquierda.

Ejemplos

1) En el conjunto $\mathcal{P}(E)$ de las partes del conjunto E , consideramos las dos leyes internas (conmutativas) $(A, B) \mapsto A \cap B$ y $(A, B) \mapsto A \cup B$. Cada una de ellas es distributiva respecto de la otra (cf. Cap. I. § 4).

2) Sea G un grupo abeliano (designado aditivamente), y designemos por $\mathcal{F}(G)$ el grupo aditivo de las aplicaciones de G en G : por definición, si $f \in \mathcal{F}(G)$ y $g \in \mathcal{F}(G)$ $f + g$ es la aplicación $x \mapsto f(x) + g(x)$.

En $\mathcal{F}(G)$ se puede definir otra ley interna, que es $(f, g) \mapsto f \circ g$ (composición de aplicaciones). La ley \circ es distributiva por la derecha respecto de la suma de $\mathcal{F}(G)$ $((f + g) \circ h = f \circ h + g \circ h)$, pero en general *no es* distributiva por la izquierda (en el caso de aplicaciones cualesquiera, en general no se verifica: $f \circ (g + h) = f \circ g + f \circ h$).

3) En el conjunto \mathbf{N}^* , la ley interna $(a, b) \mapsto ab$ (multiplicación) es distributiva respecto de la ley interna $(a, b) \mapsto \text{mcd}(a, b)$. Este lenguaje traduce la fórmula aritmética

$$c \cdot \text{mcd}(a, b) = \text{mcd}(ca, cb).$$

Análogamente, la multiplicación es distributiva respecto de la ley

$$(a, b) \mapsto \text{mcm}(a, b).$$

En general, sea $(x, y) \mapsto f(x, y)$ una ley interna sobre \mathbf{R} . Decir que la multiplicación de \mathbf{R} es distributiva respecto de esta ley, significa que, para todo $\lambda \in \mathbf{R}$, se verifica:

$$\lambda f(x, y) = f(\lambda x, \lambda y),$$

es decir que la función numérica $f: \mathbf{R}^2 \rightarrow \mathbf{R}$, es *homogénea de grado 1*.

DEFINICIÓN III.1.2

Sea Ω un dominio de operadores, E un conjunto provisto de una ley interna \top , y de una ley externa por la izquierda \perp , de dominio Ω . Diremos que **la ley \perp es distributiva respecto de \top** si para todo $\alpha \in \Omega$, y todo $x, y \in E$, se tiene: $\alpha \perp (x \top y) = (\alpha \perp x) \top (\alpha \perp y)$.

Existe una noción análoga de distributividad para leyes externas por la derecha. Supongamos que, a su vez, Ω está provisto de una ley interna \wedge : se dice que \perp es *distributiva respecto de \wedge* (o « \perp es distributiva respecto de la ley \wedge de los escalares») si

para todo $\alpha \in \Omega$, todo $\beta \in \Omega$ y todo $x \in E$, se tiene:

$$(\alpha \wedge \beta) \perp x = (\alpha \perp x) \top (\beta \perp x).$$

Ejemplos

1) Sea G un grupo abeliano (designado aditivamente). Dotamos a G de una ley externa de dominio \mathbf{Z} , por medio de la fórmula

$$(m, x) \mapsto m \cdot x \quad (m \in \mathbf{Z}, x \in G).$$

Esta ley externa es distributiva respecto de la suma de G , y respecto de la suma de los escalares.

2) Sean G, H dos grupos abelianos (designados aditivamente); sea \mathcal{A} el grupo aditivo de las aplicaciones de G en H : por definición, si

$$f \in \mathcal{A} \text{ y } g \in \mathcal{A}, \quad f + g \text{ es la aplicación } x \mapsto f(x) + g(x).$$

Designemos por $\mathcal{F}(G)$ (resp. $\mathcal{F}(H)$) el grupo de las aplicaciones de G en G (resp. de H en H). Las fórmulas

$$(\alpha, f) \mapsto \alpha \circ f \quad (\alpha \in \mathcal{F}(H), f \in \mathcal{A})$$

$$\text{y} \quad (f, \beta) \mapsto f \circ \beta \quad (\beta \in \mathcal{F}(G), f \in \mathcal{A})$$

definen sobre \mathcal{A} una ley externa por la izquierda, de dominio $\mathcal{F}(H)$, y una ley externa por la derecha, de dominio $\mathcal{F}(G)$:

La ley externa por la izquierda es distributiva respecto de los escalares (pero no respecto de la suma de \mathcal{A}) y conmuta con la composición de aplicaciones en $\mathcal{F}(H)$.

La ley externa por la derecha es distributiva respecto de la suma de \mathcal{A} (pero no respecto a los escalares) y conmuta con la composición de aplicaciones en $\mathcal{F}(G)$.

§ III.2 GENERALIDADES SOBRE LOS ANILLOS

DEFINICIÓN III.2.1

Un **anillo** es un conjunto A provisto de dos leyes de composición internas llamadas **adición** y **multiplicación** (o **suma** y **producto**) tales que

(A₁) A es un grupo abeliano para la adición.

(A₂) la multiplicación es asociativa.

(A₃) la multiplicación es distributiva respecto de la suma.

La adición de un anillo se designa, en general, con el signo $+$ (dicho de otra manera, el grupo aditivo de un anillo se designa aditivamente).

En general, la multiplicación de un anillo se designa por $(a, b) \mapsto a.b$, o $(a, b) \mapsto ab$. Para todo anillo A , escribiremos $A^* = A \setminus \{0\}$.

DEFINICIÓN III.2.2

Un anillo se llamará **unífero** ⁽¹⁾ si admite un elemento neutro para el producto, distinto de 0.

⁽¹⁾ Que sepamos, el término *unífero* sólo se emplea para las álgebras; pero un anillo con elemento unidad se puede considerar siempre como un *álgebra unífera* respecto del anillo \mathbb{Z} .

Si dicho elemento neutro existe, es único (cf. § III.2). Entonces se llama *elemento unidad* del anillo. Se le designará por e , u , I , o 1 si no se presta a confusión.

Un anillo unífero tiene, por lo menos, dos elementos: 0 y 1 .

El conjunto $\{0\}$, provisto de las leyes $0 + 0 = 0$ y $0 \cdot 0 = 0$ es un anillo, llamado *anillo nulo*. El anillo nulo no es unífero.

DEFINICIÓN III.2.3

§ Un anillo A es **conmutativo** si su multiplicación es una ley conmutativa.

Cálculo en un anillo

En un anillo cualquiera A , se tienen las siguientes propiedades (además de las que resultan de la estructura de grupo aditivo, y de la asociatividad de la multiplicación):

— Para todo $a \in A$, $a \cdot 0 = 0 = 0 \cdot a$.

En efecto: $a \cdot 0 = a(0 + 0) = a \cdot 0 + a \cdot 0$, de donde $a \cdot 0 = 0$.

Asimismo $0 \cdot a = 0$.

En particular, si A es no nulo, jamás es A un grupo respecto de la multiplicación.

— Para todo $a \in A$ y todo $b \in A$, $a(-b) = -(ab) = (-a) \cdot b$, pues $ab + a(-b) = a[b + (-b)] = a \cdot 0 = 0$, de donde $a(-b) = -(ab)$.

Análogamente $-(ab) = (-a)b$.

Estas propiedades permiten «desarrollar» los productos de sumas en un anillo. Por ejemplo:

$$(a + b)(a + b) = a^2 + ab + ba + b^2, \quad (a + b)(a - b) = a^2 - ab + ba - b^2, \quad \text{etc.}$$

Se observará que para escribir el producto de estos desarrollos se debe tener en cuenta el orden de los términos. Si A es comunicativo, se podrán, sin embargo, efectuar simplificaciones; así, las dos fórmulas anteriores se convertirán en:

$$(a + b)(a + b) = a^2 + 2ab + b^2, \quad (a + b)(a - b) = a^2 - b^2.$$

— Si A es un anillo cualquiera, para $a \in A$ y $n \in \mathbf{N}^*$, se define a^n por recurrencia:

$$a^1 = a, \quad a^n = a^{n-1} \cdot a.$$

La asociatividad del producto demuestra que para $m, n \in \mathbf{N}^*$, se tiene: $a^{m+n} = a^m \cdot a^n$.

Cuando el anillo A es unífero, se define además $a^0 = 1$, y la fórmula anterior es válida para $m, n \in \mathbf{N}$.

Vamos a precisar estas reglas de cálculo en un anillo conmutativo.

Recordemos que se ha definido ya, para todo $m \in \mathbf{Z}$ y todo $x \in A$, el símbolo $m \cdot x$ (cf. Cap. II).

TEOREMA III.2.1

Sea A un anillo **conmutativo**. Para todos los elementos a, b de A y todo entero $n \geq 1$, se tiene la fórmula

$$(1) \quad (a + b)^n = a^n + \left(\sum_{p=1}^{n-1} \binom{n}{p} \cdot a^{n-p} \cdot b^p \right) + b^n$$

(fórmula del binomio).

Demostración. Por recurrencia sobre n . La fórmula es evidente para $n = 1$. Supongámosla verdadera para el entero n , y probemos que es verdadera para el entero $n + 1$; se tiene, en efecto

$$\begin{aligned} (a + b)^{n+1} &= (a + b)^n (a + b) = \left(a^n + \sum_{p=1}^{n-1} \binom{n}{p} \cdot a^{n-p} b^p + b^n \right) (a + b) \\ &= a^{n+1} + b^{n+1} + a^n b + ab^n + \sum_{p=1}^{n-1} \binom{n}{p} a^{n+1-p} b^p + \\ &\quad + \sum_{p=1}^{n-1} \binom{n}{p} a^{n-p} b^{p+1} \end{aligned}$$

(puesto que el producto es asociativo y conmutativo).

Pero podemos escribir:

$$\begin{aligned} \sum_{p=1}^{n-1} \binom{n}{p} a^{n+1-p} b^p &= n \cdot a^n b + \sum_{p=2}^{n-1} \binom{n}{p} a^{n+1-p} b^p, \\ \sum_{p=1}^{n-1} \binom{n}{p} a^{n-p} b^{p+1} &= n \cdot ab^n + \sum_{k=2}^{n-1} \binom{n}{k-1} a^{n+1-k} b^k, \end{aligned}$$

de donde

$$\begin{aligned} (a + b)^{n+1} &= a^{n+1} + b^{n+1} + (n + 1) a^n \cdot b + (n + 1) \cdot ab^n + \\ &\quad + \sum_{p=2}^{n-1} \left(\binom{n}{p} + \binom{n}{p-1} \right) a^{n+1-p} b^p \end{aligned}$$

se concluye con la ayuda de las fórmulas

$$\binom{n}{p} + \binom{n}{p-1} = \binom{n+1}{p}, \quad \text{y} \quad \binom{n+1}{1} = \binom{n+1}{n} = n + 1 \text{ .c.q.d.}$$

Si A es unífero (y conmutativo) la fórmula del binomio se puede escribir

$$(a + b)^n = \sum_{p=0}^n \binom{n}{p} a^{n-p} b^p = \sum_{\substack{p \geq 0, q \geq 0 \\ p+q=n}} \frac{n!}{p! q!} a^p b^q.$$

Vamos a generalizar la fórmula del binomio en el caso de un anillo *unífero* conmutativo;

TEOREMA III.2.2

Sea A un anillo unífero conmutativo, y sean a_1, a_2, \dots, a_n elementos de A .
Para todo entero m , se tiene:

$$(2) \quad (a_1 + a_2 + \dots + a_n)^m = \sum_{\substack{p_1, p_2, \dots, p_n \geq 0 \\ p_1 + p_2 + \dots + p_n = m}} \frac{m!}{p_1! p_2! \dots p_n!} a_1^{p_1} a_2^{p_2} \dots a_n^{p_n}.$$

(Esta fórmula se denomina a veces «fórmula del binomio, generalizada».)*

Demostración. Por recurrencia sobre n . La fórmula es verdadera para $n = 2$ según el teorema III.2.1. Supongámosla verdadera para el entero n y probemos que es verdadera para $n + 1$; se tiene:

$$(a_1 + a_2 + \dots + a_n + a_{n+1})^m = (a + a_{n+1})^m = \sum_{\substack{p+q=m \\ p, q \geq 0}} \frac{m!}{p! q!} a^p a_{n+1}^q,$$

si hacemos $a = a_1 + a_2 + \dots + a_n$. Luego, por hipótesis,

$$a^p = \sum_{\substack{p_1, p_2, \dots, p_n \geq 0 \\ p_1 + p_2 + \dots + p_n = p}} \frac{p!}{p_1! p_2! \dots p_n!} a_1^{p_1} \dots a_n^{p_n}.$$

De donde,

$$(a + a_{n+1})^m = \sum_{\substack{p_1, p_2, \dots, p_n, q \geq 0 \\ p_1 + p_2 + \dots + p_n + q = m \\ p_1 + p_2 + \dots + p_n = p}} \frac{m!}{p! q!} \cdot \frac{p!}{p_1! p_2! \dots p_n!} a_1^{p_1} \dots a_n^{p_n} a_{n+1}^q,$$

y la fórmula buscada se obtiene haciendo $q = p_{n+1}$. c.q.d.

* En el texto original a esta fórmula se la llama «fórmula del multinomio» pero hemos preferido denominarla fórmula del binomio, generalizada. (N. del T.).

Notas

- 1) Según el capítulo I, § 10, el número de términos del segundo miembro de
 (2) es $\binom{m+n-1}{n-1}$.
 2) En particular se observará la fórmula:
 $(a+b+c)^3 = a^3 + b^3 + c^3 + 3(a^2b + a^2c + b^2c + b^2a + c^2a + c^2b) + 6abc$.

TEOREMA III.2.3

En un anillo conmutativo, se tienen las fórmulas siguientes, válidas para $a \in A$, $b \in A$:

$$(3) \quad \begin{cases} a^2 - b^2 = (a - b)(a + b), \\ a^n - b^n = (a - b) \left(a^{n-1} + b^{n-1} + \sum_{k=1}^{n-2} a^{n-1-k} b^k \right). \end{cases}$$

Si A es, además, unífero, las fórmulas (3) se pueden condensar en la fórmula única

$$(4) \quad a^n - b^n = (a - b) \left(\sum_{k=0}^{n-1} a^{n-1-k} b^k \right).$$

Demostración. Cuando A es unífero, (4) resulta de (3). Demostremos la segunda fórmula (3). Si designamos por D el miembro de la derecha, podemos escribir:

$$\begin{aligned} D &= a^n + ba^{n-1} + \sum_{k=1}^{n-2} a^{n-k} b^k - a^{n-1} b - b^n - \sum_{k=1}^{n-2} a^{n-1-k} b^{k+1} \\ &= a^n - b^n - a^{n-1} b + b^{n-1} a - \left(\sum_{k=2}^{n-1} a^{n-k} b^k \right) + \left(\sum_{k=2}^{n-2} a^{n-k} b^k \right) + a^{n-1} b \\ &= a^n - b^n. \text{ c.q.d.} \end{aligned}$$

Como caso particular, se tendrá presente la fórmula que sigue, válida también en un anillo unífero *cualquiera* (no necesariamente conmutativo)

$$a^n - 1 = (a - 1)(1 + a + a^2 + \cdots + a^{n-1}).$$

Nota importante. Sea A un anillo (resp. un anillo unífero) no necesariamente conmutativo, y $a \in A$, $b \in A$. Entonces si $ab = ba$ la relación (3) (resp. (4)) es válida. Para verlo, basta reemplazar A por el menor subanillo B de A que

contenga a a y b (resp. por el menor subanillo unífero B de A que contenga a a y b): en efecto, la relación $ab = ba$ implica que B es conmutativo, y III.2.3 se cumple.

Elementos invertibles

DEFINICIÓN III.2.4

Sea A un anillo unífero, y $a \in A$. Por definición:

- a es **invertible por la izquierda** si existe $b \in A$ tal que $ba = 1$, a se le llama un inverso de a por la izquierda.
- a es **invertible por la derecha** si existe $c \in A$ tal que $ac = 1$, a se le llama un inverso de a por la derecha.
- a es **invertible** si es invertible por la derecha y por la izquierda.

A un elemento invertible se le llama a veces una *unidad** del anillo. No se debe confundir esta noción de unidad con el *elemento unidad* del anillo, que es único y se designa por 1.

Si a es invertible por la izquierda, es *regular por la izquierda*, pues (si b es un inverso por la izquierda de a), la igualdad $ax = ay$ implica

$$b(ax) = b(ay) = 1.x = 1.y = x = y.$$

Análogamente, todo elemento invertible por la derecha es *regular por la derecha*.

Los recíprocos son falsos (cf. ejemplos dados más abajo).

Si a es invertible por la izquierda (resp. por la derecha) el inverso por la izquierda (resp. por la derecha) no es necesariamente único (cf. ejemplos).

Si a es invertible, el inverso por la derecha y el inverso por la izquierda coinciden, y dicho inverso es único, pues, de $ac = 1$ y $ba = 1$, se deduce:

$$bac = (ba)c = b(ac) = c = b,$$

y, de $ba = b'a = 1$, se deduce (suponiendo siempre que $ac = 1$):

$$bac = b'ac = b = b' = c.$$

El elemento unidad del anillo es evidentemente invertible.

Si A es conmutativo, la existencia de un inverso por la derecha para $a \in A$ equivale a la de un inverso por la izquierda, y por lo tanto a la de un inverso único de a en A .

(*) Ciertos autores llaman a los elementos unidad de un anillo elementos *unitarios*. (N. del T.).

TEOREMA III.2.4

Sea A un anillo unífero, y sea $U(A)$ el conjunto de los elementos invertibles de A . Entonces $U(A)$ es estable por la multiplicación de A , y la ley inducida sobre $U(A)$ por la multiplicación de A es una ley de grupo, cuyo elemento neutro es 1.

Demostración. La única cosa que se ha de probar es la estabilidad de A para el producto, pues el resto resultará de las consideraciones que preceden a este teorema. Sean $u \in U(A)$ y $v \in U(A)$, y designemos por u' y v' a los inversos de u y v . Se tiene:

$$(v' u') (uv) = v'(u' u) v = v' 1 v = v' v = 1,$$

y

$$(uv) (v' u') = u(vv') u' = u \cdot 1 \cdot u' = uu' = 1,$$

luego uv admite a $v'u'$ como inverso por la derecha y por la izquierda, de donde $u \cdot v \in U(A)$. c.q.d.

$U(A)$ es el llamado *grupo de unidades* de A . Su conocimiento da, a menudo, información estimable acerca de la estructura de A .

Divisores de cero

Sea A un anillo no nulo. Un elemento $a \in A$ es *divisor de 0 por la izquierda* si $a \neq 0$, y si existe $b \in A$, $b \neq 0$, tal que $ab = 0$. Análogamente, a es *divisor de 0 por la derecha*, si existe $c \neq 0$ tal que $ca = 0$ ($a \neq 0$).

Decir que a es un divisor de 0 por la derecha equivale a decir que a no es regular por la derecha: pues si $ba \neq 0$ con $b \neq 0$, a no es regular por la derecha puesto que $0 \cdot a = 0$. Y si $a \cdot b = ac$ con $b \neq c$, se deduce

$$a(b - c) = 0 \quad \text{con} \quad b - c \neq 0.$$

Análogamente, a es divisor de cero por la izquierda si, y sólo si, a no es regular por la izquierda.

Si A es conmutativo, las nociones de divisor de cero por la derecha y por la izquierda coinciden.

Nota. En un anillo pueden existir, en general, divisores de cero por la izquierda (resp. por la derecha) que sean regulares por la derecha (resp. por la izquierda).

DEFINICIÓN III.2.5

Sea A un anillo no nulo; se dice que A es **íntegro** si no admite divisores de cero.

Un anillo íntegro y conmutativo se llama, a veces, *dominio de integridad*. En un anillo íntegro, la relación $ab = 0$ es, pues, equivalente a la relación:

$$(a = 0 \quad \text{o} \quad b = 0);$$

y todos los elementos $\neq 0$ son regulares.

Ejemplos

1) El conjunto \mathbf{Z} de los enteros relativos, dotado de la adición y de la multiplicación ordinarias, es un anillo unífero, conmutativo e íntegro. El grupo de las unidades (elementos invertibles) es $\{-1, 1\}$.

2) Sea n un entero > 0 , y sea χ el homomorfismo canónico de grupos aditivos: $\chi: \mathbf{Z} \rightarrow \mathbf{Z}/n\mathbf{Z}$. Se verifica que $(\chi(a) = \chi(a') \text{ y } \chi(b) = \chi(b'))$ implican $\chi(ab) = \chi(a'b')$; luego $\chi(ab)$ sólo depende de $\chi(a)$ y $\chi(b)$, y, haciendo $\chi(a) \cdot \chi(b) = \chi(ab)$, se define una ley interna sobre $\mathbf{Z}/n\mathbf{Z}$; de este modo, $\mathbf{Z}/n\mathbf{Z}$ se convierte en un anillo unífero y conmutativo, de elemento unidad $\chi(1)$, para la adición habitual y para el producto, llamado *anillo de las clases módulo n* . Como ejemplo, construimos las tablas de sumar y de multiplicar de $\mathbf{Z}/4\mathbf{Z}$ (\bar{u} designa a $\chi(u)$):

Adición					Multiplicación				
	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$		$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Vemos que el grupo de las unidades de $\mathbf{Z}/4\mathbf{Z}$ es $\{\chi(1), \chi(3)\}$.

En el caso general (n cualquiera) la relación $\chi(ab) = \chi(0)$ equivale a:
 $ab \equiv 0 \pmod{n}$ (es decir, ab es divisible por n).

Se deduce que $\mathbf{Z}/n\mathbf{Z}$ es íntegro si, y sólo si, n es un número primo.

Sean m y a dos enteros; en virtud de la definición del producto de $\mathbf{Z}/n\mathbf{Z}$, se tiene:

$$m \cdot \chi(a) = \chi(ma).$$

Luego $\chi(a)$ es invertible si, y sólo si, existe un entero $m \in \mathbf{Z}$ tal que $m \cdot \chi(a) = \chi(1)$; lo que significa que $\chi(a)$ es un generador del grupo $\mathbf{Z}/n\mathbf{Z}$. Así pues, sabemos (§ II.5) que estos generadores son los $\chi(a)$ tales que

$$0 \leq a \leq n-1, \quad \text{y} \quad (a, n) = 1.$$

Luego los elementos invertibles de $\mathbf{Z}/_n\mathbf{Z}$ son las clases $\chi(a)$, en donde a es un entero $< n$ y primo con n .

3) Si A es un anillo unífero y E un conjunto, el conjunto $\mathcal{F}(E, A)$ de las aplicaciones de E en A se convierte en un anillo unífero si se le dota de las leyes siguientes (E se supone no vacío):

adición $(f, g) \mapsto f + g$ definida por $(f + g)(x) = f(x) + g(x) \ (x \in E)$;

multiplicación $(f, g) \mapsto fg$ definida por $(fg)(x) = f(x) \cdot g(x) \ (x \in E)$.

El elemento unidad de $\mathcal{F}(E, A)$ es la aplicación constante igual a 1.

Si A es conmutativo, $\mathcal{F}(E, A)$ es conmutativo.

4) Si G es un grupo abeliano no reducido a $\{0\}$ (designado aditivamente), sea $\mathcal{H}(G)$ el grupo aditivo de los endomorfismos de G . Dotamos a $\mathcal{H}(G)$ de la ley interna

$$(f, g) \mapsto f \circ g.$$

$\mathcal{H}(G)$ se convierte entonces en un anillo unífero. El hecho de que f, g, h sean endomorfismos de G nos asegura las fórmulas de distributividad

$$f \circ (g + h) = f \circ g + f \circ h, \quad (f + g) \circ h = f \circ h + g \circ h,$$

el elemento unidad de $\mathcal{H}(G)$ es la aplicación idéntica de G . En general, $\mathcal{H}(G)$ no es conmutativo ni íntegro.

5) *Anillo de Boole.* Designamos por A al anillo $\mathbf{Z}/_2\mathbf{Z}$. Para todo conjunto E , la aplicación $f \mapsto f^{-1}(\chi(1))$ (en donde f designa a una aplicación de E en A) es una biyección de $\mathcal{F}(E, A)$ en el conjunto $\mathcal{P}(E)$ de las partes de E . Cuando se transporta, con la ayuda de esta biyección, la estructura de anillo de $\mathcal{F}(E, A)$ sobre $\mathcal{P}(E)$, se obtiene en $\mathcal{P}(E)$ una estructura de anillo unífero conmutativo; la adición y la multiplicación de $\mathcal{P}(E)$ están definidas por

$$X + Y = (X \cup Y) \setminus (X \cap Y)$$

$$X \cdot Y = X \cap Y.$$

El elemento unidad es E , el elemento nulo es \emptyset ; para todo $X \in \mathcal{P}(E)$, se tiene

$$X + X = 0 \quad \text{y} \quad X^2 = X.$$

Estas propiedades se expresan diciendo que $\mathcal{P}(E)$ es un *anillo de Boole*. El grupo de unidades es $\{E\}$. (En el bien entendido que E se suponga no vacío.)

* 6) Sea G un grupo abeliano no reducido a $\{0\}$ y $P = G^{\mathbb{N}}$ el grupo producto, formado con las sucesiones $(a_n)_{n \geq 0}$ de elementos de G ; designemos por $\mathcal{F}(P)$ el anillo de los endomorfismos de P . Para toda sucesión $a = (a_n)_{n \geq 0}$ hacemos:

$$f(a) = (b_n)_{n \geq 0}, \quad \text{con } b_0 = 0, \quad \text{y } b_n = a_{n-1} \quad \text{para } n \geq 1;$$

$$g(a) = (c_n)_{n \geq 0}, \quad c_n = a_{n+1} \quad \text{para } n \geq 0;$$

$$h(a) = (d_n)_{n \geq 0}, \quad \text{con } d_0 = a_0, \quad \text{y } d_n = 0 \quad \text{para } n \geq 1.$$

f, g, h son elementos de $\mathcal{F}(P)$, $f \neq 0, g \neq 0, h \neq 0$. Designemos por I al elemento unidad de $\mathcal{F}(P)$ (aplicación idéntica de P). Se tiene:

$$g \circ f = I, \text{ luego } f \text{ es invertible por la izquierda.}$$

Además: $h \circ f = 0$, luego f es un divisor de cero por la izquierda.

Finalmente: $(g + h) \circ f = g \circ f + h \circ f = I$, luego f admite dos inversos por la izquierda distintos: g y $g + h$.

§ III.3 SUBANILLOS Y ANILLOS PRODUCTOS

Según las definiciones generales del capítulo II, se puede establecer:

DEFINICIÓN III.3.1

$\left\{ \begin{array}{l} \text{Sea } A \text{ un anillo; una parte } B \text{ de } A \text{ es un } \mathbf{subanillo} \text{ de } A \text{ si } B \text{ es un} \\ \text{subgrupo del grupo aditivo de } A, \text{ y si } B \text{ es estable para la multiplicación de } A. \end{array} \right.$

Cuando así ocurre, las leyes inducidas en B por las de A hacen de B un anillo. Es evidente que si A es conmutativo, B es conmutativo, y que si A es íntegro, B es íntegro. Por el contrario, si A es unífero, B no es necesariamente unífero: por ejemplo, $2\mathbb{Z}$ es un subanillo no unífero de \mathbb{Z} .

Además, si A y B son uníferos, puede ocurrir que los elementos unidad de A y B sean distintos.

Ejemplo (cf. Cap. VIII)

Sea A el anillo $M_2(K)$ de las matrices cuadradas de orden 2 sobre un cuerpo conmutativo K . A es unífero de elemento unidad

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (1, \text{ elemento unidad de } K).$$

Sea B el subanillo formado por las matrices de la forma:

$$\begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}, \quad \lambda \in K.$$

B es un subanillo unífero de A , de elemento unidad

$$J = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{y} \quad J \neq I. \text{ c.q.d.}$$

Sea B un subanillo del anillo A ; supongamos que A y B son uníferos y designemos por I (resp. J) al elemento unidad de A (resp. de B). Si $I \neq J$, J es un divisor de cero en A , pues se tiene:

$$IJ = J = J^2, \text{ de donde } J \cdot (I - J) = 0 = (I - J) \cdot J.$$

● Si A es un anillo unífero, reservaremos el término *subanillo unífero de A* para designar a un subanillo unífero que posea el mismo elemento unidad que A .

DEFINICIÓN III.3.2

Sean A y B dos anillos. Se llama **anillo producto** de los anillos A y B , y se designa $A \times B$, al conjunto $A \times B$ provisto de las leyes definidas por:

$$(a, b) + (a', b') = (a + a', b + b')$$

$$(a, b) \cdot (a', b') = (aa', bb') \quad a, a' \in A; b, b' \in B.$$

Se verifica inmediatamente que estas leyes definen en $A \times B$ una estructura de anillo.

Si A y B son uníferos, $A \times B$ es unífero, de elemento unidad $(1, 1)$.

Si A y B son conmutativos, $A \times B$ es conmutativo.

Si A y B son no nulos, $A \times B$ no es íntegro, pues

$$(a, 0) \cdot (0, b) = (0, 0) = 0.$$

Sean A y B dos anillos uníferos, y $C = A \times B$. Sea $c \in C$, $c = (a, b)$; c es invertible por la izquierda (resp. por la derecha) si, y sólo si, existe $u \in A$ y $v \in B$ tales que

$$(ua, vb) = (1, 1) \quad (\text{resp. } (au, bv) = (1, 1)).$$

En particular c es invertible si, y sólo si, a y b son invertibles en A y B . Se deduce sin dificultad el resultado siguiente:

Si $U(A)$, $U(B)$, $U(C)$ designan respectivamente a los grupos de unidades de los anillos uníferos A , B , C , en donde $C = A \times B$, se tiene: $U(C) = U(A) \times U(B)$.

En general, sea $(A_i)_{i \in I}$ una familia de anillos; el conjunto producto

$$A = \prod_{i \in I} A_i$$

dotado de la adición producto de las operaciones de adición de los A_i , y de la multiplicación siguiente:

$$\text{si } x = (x_i)_{i \in I} \text{ e } y = (y_i)_{i \in I}, \quad x \cdot y = (x_i y_i)_{i \in I} \text{ (multiplicación producto),}$$

es un anillo, llamado *anillo producto de los A_i* . Si todos los A_i son uníferos, A es unífero, y si todos los A_i son conmutativos, A es conmutativo.

Si todos los A_i son iguales a un mismo anillo A , el anillo $\prod_{i \in I} A_i$ (designado por A^I) se puede identificar con el anillo de las aplicaciones de I en A , o sea a $\mathcal{F}(I, A)$.

§ III.4 HOMOMORFISMOS, IDEALES, Y ANILLOS COCIENTES

Conforme a las definiciones generales del capítulo II, se puede establecer:

DEFINICIÓN III.4.1

Sean A , B dos anillos. Una aplicación $\rho : A \rightarrow B$ es un homomorfismo (de anillos) si ρ es un homomorfismo para los grupos aditivos de A y B , y si, para todo $a \in A$ y todo $b \in A$, se tiene:

$$\rho(ab) = \rho(a) \rho(b).$$

Es claro que si $\rho : A \rightarrow B$ es un homomorfismo de anillos, $\rho(A)$ es un subanillo de B .

Si A y B son uníferos, un homomorfismo $\rho : A \rightarrow B$ no tiene porque transformar necesariamente el elemento unidad de A en el elemento unidad de B . Por ejemplo, si para todo $x \in A$, $\zeta(x) = 0$, ζ es un homomorfismo (llamado *homomorfismo nulo*, y designado por 0). Otro ejemplo es el siguiente: A y B son uníferos, C es el anillo $A \times B$, $\rho : A \rightarrow C$ está definido por $\rho(a) = (a, 0)$; ρ es un homomorfismo, y $\rho(1) = (1, 0) \neq (1, 1)$.

Convenio

● En todo lo que sigue a lo largo de esta obra, si A y B designan anillos uníferos se llamará «homomorfismo de anillos de A en B » tanto al homomorfismo nulo como a un homomorfismo que transforme el elemento unidad de A en el de B . A los restantes homomorfismos se les llamará *representaciones de A en B* .

Ideales

● Para no recargar el texto, el estudio de los ideales y de los anillos cocientes lo realizaremos sólo para *anillos conmutativos uníferos*.

DEFINICIÓN III.4.2

} Una parte α del anillo **unífero conmutativo** A es un **ideal** si es
 } un subgrupo del grupo aditivo de A , y si, para todo $a \in \alpha$ y todo $\lambda \in A$,
 } se tiene: $\lambda a \in \alpha$.

La segunda condición implica la relación:

$$[(a \in \alpha) \Rightarrow (-a = -1 \cdot a \in \alpha)] ;$$

por consiguiente, para que la parte α de A sea un ideal de A , es necesario y suficiente que se verifiquen las dos condiciones siguientes:

- (1) $[(a \in \alpha) \text{ y } (b \in \alpha)] \Rightarrow [a + b \in \alpha] ;$
 (2) $[(a \in \alpha) \text{ y } (\lambda \in A)] \Rightarrow [\lambda a \in \alpha] .$

Según (2), si $1 \in \alpha$, $\alpha = A$; recíprocamente, si $\alpha = A$, $1 \in \alpha$; enunciamos:

III.4.1 Para que el ideal α del anillo A sea igual a A , es necesario y suficiente
 || que: $1 \in \alpha$.

Propiedades inmediatas de los ideales

$\{0\}$ y A son ideales (a los ideales distintos de A se les llama *ideales propios*).
 Todo ideal contiene al $\{0\}$.

Si $(\alpha_i)_{i \in I}$ es una familia de ideales de A , $\bigcap_{i \in I} \alpha_i$ es un ideal. Es el ínfimo de los α_i en $\mathcal{P}(A)$ ordenado por inclusión. Esta propiedad nos conducirá a la definición III.4.3.

Sea $\rho : A \rightarrow B$ un homomorfismo de anillos. Para todo ideal \mathfrak{b} de B , $\rho^{-1}(\mathfrak{b})$ es un ideal de A . En particular, el **núcleo** $\rho^{-1}(0)$ es un ideal de A . Por el contrario, si α es un ideal de A , $\rho(\alpha)$ no es necesariamente un ideal de B , salvo que ρ sea epiyectivo (en efecto, en el caso general, (2) no se verifica para $\rho(\alpha)$; y $\rho(\alpha)$ verifica (2) siempre que ρ es epiyectivo).

DEFINICIÓN III.4.3

Sea S una parte del anillo A . A la intersección de la familia de los ideales de A que contienen a S (que es un ideal en virtud de lo que precede) se denomina **ideal engendrado por S** ; a S se le llama el **sistema de generadores** de este ideal.

Si S es no vacío, el ideal engendrado por S contiene todos los elementos de la forma

$$\lambda_1 s_1 + \lambda_2 s_2 + \dots + \lambda_n s_n \quad (\lambda_1, \dots, \lambda_n \in A ; s_1, \dots, s_n \in S) , \text{ según (1) y (2).}$$

Recíprocamente, el conjunto de los elementos de la forma anterior es un ideal ((2) se verifica inmediatamente; (1) es un poco más delicado: ver § 8 para un razonamiento general).

Enunciaremos:

III.4.2 El ideal engendrado por una parte no vacía S del anillo A es igual al conjunto de los elementos de la forma $\lambda_1 s_1 + \lambda_2 s_2 + \dots + \lambda_n s_n$, en donde los λ_i son elementos cualesquiera de A , y los s_i elementos cualesquiera de S .

Casos particulares

1) Si S es una parte finita $\{a_1, \dots, a_n\}$ de A , el ideal engendrado por S es el conjunto de elementos de la forma $\sum_{i=1}^n \lambda_i a_i$ ($\lambda_i \in A$). Este ideal se designa por «ideal (a_1, \dots, a_n) » o, más simplemente, por (a_1, a_2, \dots, a_n) cuando no se preste a confusión.

● A un ideal de la forma (a) , es decir un ideal engendrado por un elemento a , le llama **ideal principal**; a es un **generador** del ideal principal (a) . Un ideal principal puede admitir generadores distintos.

2) Si $S = \bigcup_{i \in I} \alpha_i$, en donde $(\alpha_i)_{i \in I}$ es una familia no vacía de ideales de A , el ideal engendrado por S coincide con el conjunto de los elementos de la forma $\sum_{i \in J} a_i$, en donde J es una parte finita cualquiera de I , y para todo i , $a_i \in \alpha_i$, en virtud de III.4.2 y de la condición (2).

Por esta razón, al ideal engendrado por $\bigcup_{i \in I} \alpha_i$ se le llama *suma* de los ideales α_i , y se designa por $\sum_{i \in I} \alpha_i$. En el conjunto de ideales de A ordenado por inclusión, $\sum_{i \in I} \alpha_i$ es el *supremo* de los ideales α_i . En general, este ideal *no es igual* a $\bigcup_{i \in I} \alpha_i$ (cf. Cap. II, notas que siguen a la definición II.4.3).

En virtud de su importancia, resumimos en forma de teorema las propiedades que acabamos de ver, relativas al supremo e ínfimo de una familia de ideales:

TEOREMA III.4.3

|| En el conjunto (ordenado por inclusión) de los ideales de un anillo (conmutativo, unífero), toda familia de ideales $(\alpha_i)_{i \in I}$ admite un supremo y un ínfimo. El supremo es $\sum_{i \in I} \alpha_i$ (ideal engendrado por $\bigcup_{i \in I} \alpha_i$), el ínfimo es $\bigcap_{i \in I} \alpha_i$.

Anillos cocientes

Sea α un ideal del anillo conmutativo A ; designemos por $p : A \rightarrow A/\alpha$ al homomorfismo canónico del grupo aditivo A en el grupo cociente A/α (cf. Cap. II, § 5).

En primer lugar vemos que, para todo $a \in A$ y todo $b \in B$, el elemento $p(ab)$ depende sólo de $p(a)$ y de $p(b)$ (en otras palabras, la relación de equivalencia $x \sim y \Leftrightarrow x - y \in \alpha$ es compatible con la multiplicación de A). En efecto, si $a' \in p(a)$ y $b' \in p(b)$, se tiene

$$a' = a + u, \quad b' = b + v, \quad \text{con } u \in \alpha, v \in \alpha,$$

de donde

$$a' b' = ab + av + bu + uv.$$

Puesto que α es un ideal, $bu \in \alpha$, $av \in \alpha$ y $uv \in \alpha$, se deduce:

$$p(a' b') = p(ab),$$

según se había enunciado.

Podemos, pues, definir en A/α una ley de composición interna, escribiendo para $X \in A/\alpha$ e $Y \in A/\alpha$: $XY = p(xy)$ (x es un elemento cualquiera de X e y es un elemento cualquiera de Y). Si $\alpha = A$, A/α es el anillo nulo, y en caso contrario no se comprueba fácilmente que la multiplicación así definida convierte a A/α en un

anillo unífero, conmutativo, de elemento unidad $I = p(1)$. Como ejemplo, comprobamos la distributividad. Puesto que A es conmutativo, bastará comprobar la fórmula:

$$Z(X + Y) = ZX + ZY \quad (X, Y, Z \in A/\alpha);$$

Para ello se eligen $x \in X$, $y \in Y$, $z \in Z$, y se tiene sucesivamente:

$$\begin{aligned} X + Y &= p(x + y), \quad Z = p(z), \quad Z(X + Y) = p(z(x + y)) = p(zx + zy) \\ &= p(zx) + p(zy) = ZX + ZY. \end{aligned}$$

DEFINICIÓN III.4.4

Sea α un ideal de A ; al anillo A/α definido antes se le llama **anillo cociente** de A por α . Al homomorfismo de anillos

$$p : A \rightarrow A/\alpha$$

se le llama **homomorfismo canónico**.

A/A es el anillo nulo, y $A/\{0\} = A$.

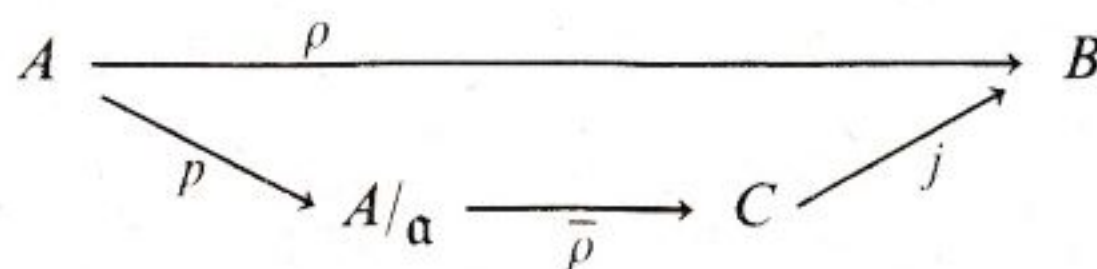
TEOREMA III.4.4 (descomposición canónica de un homomorfismo)

Sean A y B dos anillos conmutativos uníferos, y $\rho : A \rightarrow B$ un homomorfismo de anillos. Designemos por α al núcleo de ρ , por C a la imagen de ρ , por $p : A \rightarrow A/\alpha$ a la aplicación canónica y por $j : C \rightarrow B$ a la inyección canónica; sea $\bar{\rho} : A/\alpha \rightarrow C$ el isomorfismo de grupos tal que

$$(1) \quad \rho = j \circ \bar{\rho} \circ p.$$

Entonces j es un homomorfismo de anillos, y $\bar{\rho}$ es un isomorfismo de anillos.

Demostración



Según el capítulo II, § 5, sabemos que p , j y $\bar{\rho}$ son homomorfismos de grupos, y que $\bar{\rho}$ es el único homomorfismo de grupos que verifica (1). Hemos visto antes que p es un homomorfismo de anillos. Todo consiste, pues, en demostrar que j y $\bar{\rho}$ son homomorfismos de anillos. Es evidente si $\alpha = A$. Supongamos que $\alpha \neq A$;

puesto que $\rho(1) = 1$, C es entonces un subanillo unífero de B , y j es realmente un homomorfismo. Finalmente sea

$$X \in A/\alpha \text{ y } Y \in A/\alpha, \quad x \in X \text{ y } y \in Y;$$

se tiene:

$$\begin{aligned}\bar{\rho}(XY) &= \bar{\rho}[p(xy)] = \rho(xy) = \rho(x) \rho(y) \\ &= (\bar{\rho}[p(x)]) (\bar{\rho}[p(y)]) = \bar{\rho}(X) \bar{\rho}(Y),\end{aligned}$$

que demuestra que $\bar{\rho}$ es realmente un homomorfismo de anillos, si se tiene en cuenta que la unidad $I = p(1)$ de A/α verifica $\bar{\rho}(I) = \rho(1) = 1$. c.q.d.

Nota. Sea A un anillo unífero conmutativo, y B un anillo unífero *cualquiera*, y sea $\rho: A \rightarrow B$ un homomorfismo no nulo. $\rho(A)$ es un subanillo *conmutativo* unífero de B , y por lo tanto podemos aplicar III.4.1 a ρ , considerándolo una aplicación de A en $\rho(A)$. Designando por α a su núcleo (que es el núcleo de ρ), vemos que A/α es, también, isomorfo a $\rho(A)$.

El teorema III.4.4 es de vital importancia, y presta los mismos servicios que el teorema análogo de la teoría de grupos.

* Aplicaciones

1) Sabemos que los únicos subgrupos de \mathbf{Z} son los conjuntos $n\mathbf{Z}$, n entero ≥ 0 . Puesto que, evidentemente, estos conjuntos son ideales, son los únicos ideales. Luego todo anillo cociente de \mathbf{Z} es un anillo $\mathbf{Z}/n\mathbf{Z}$ para un valor conveniente de n . Sea ahora A un anillo unífero *cualquiera*, de elemento unidad I .

La aplicación $\varphi: \mathbf{Z} \rightarrow A$, tal que $\varphi(m) = m \cdot I$, es un homomorfismo de anillos (a causa de la distributividad del producto con respecto de la suma en A).

El núcleo de φ es de la forma $q\mathbf{Z}$, para un entero $q \in \mathbf{N}$.

Según el teorema III.4.4, la imagen P de φ es un subanillo de A isomorfo a $\mathbf{Z}/q\mathbf{Z}$ (por lo tanto a \mathbf{Z} si $q = 0$), luego P es el menor subanillo unífero de A . Estas consideraciones nos conducen a la siguiente:

DEFINICIÓN III.4.5

La característica de un anillo unífero A , de elemento unidad I , es el entero $q \geq 0$ tal que $q\mathbf{Z}$ sea el núcleo del homomorfismo

$$m \mapsto m \cdot I$$

de \mathbf{Z} en A ⁽¹⁾.

⁽¹⁾ Para los especialistas en álgebra conmutativa, esta definición de la característica se presta a veces a confusión, y por ello se reserva preferentemente para *cuerpos conmutativos*. Sin embargo, en el nivel en que nos movemos, no hay inconveniente alguno en utilizar esta definición.

Cuando la característica es nula, $m \mapsto m.I$ es un isomorfismo de \mathbf{Z} en un subanillo de A (designado antes por P). En este caso, se puede decir que A «contiene» a \mathbf{Z} . Cuando la característica q es > 0 , q es el menor entero $m > 0$ tal que $m.I = 0$, y todo entero m tal que $m.I = 0$ es un múltiplo de q . Además, para todo $x \in A$, $q.x = (q.I)x = 0$.

Cuando A es íntegro y la característica q de A es > 0 , el anillo P es íntegro (todo subanillo de un anillo íntegro es íntegro). Dado que P y $\mathbf{Z}/q\mathbf{Z}$ son isomorfos, $\mathbf{Z}/q\mathbf{Z}$ es íntegro, luego q es primo. Enunciaremos:

III.4.5 *Un anillo íntegro y de característica no nula tiene por característica un número primo.*
Por otra parte es evidente que todo anillo finito tiene una característica no nula.

Nota. Supongamos que A es un anillo conmutativo cuya característica es un número primo p . Para todos $x, y \in A$, se tiene:

$$(3) \quad (x + y)^p = x^p + y^p.$$

En efecto, la fórmula del binomio proporciona:

$$(x + y)^p = x^p + y^p + \sum_{k=1}^{p-1} \binom{p}{k} x^k y^{p-k},$$

y para $1 \leq k \leq p-1$, se tiene:

$$\binom{p}{k} \equiv 0 \pmod{p}. \text{ En efecto } \binom{p}{k} = p \frac{(p-1) \dots (p-k+1)}{k!} = p \cdot \frac{m}{k!}$$

en donde m es un entero. $k!$ es un producto de números primos con p , luego es primo con p , y puesto que $k!$ divide a pm , $k!$ debe dividir a m según el teorema de Gauss, y de ahí nuestra afirmación.

Se deduce, para $1 \leq k \leq p-1$, $\binom{p}{k} x^k y^{p-k} = 0$, y la fórmula del binomio se reduce a (3).

En general, se deduce, por recurrencia sobre el entero $\alpha \geq 1$:

$$(x + y)^{p^\alpha} = x^{p^\alpha} + y^{p^\alpha}.$$

2) *Fórmula de Euler*

Sean m y n dos enteros > 0 y primos entre sí. Definimos la aplicación

$$f: \mathbf{Z} \rightarrow \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}, \quad \text{por: } f(x) = (p(x), q(x)),$$

en donde $p: \mathbf{Z} \rightarrow \mathbf{Z}/m\mathbf{Z}$ y $q: \mathbf{Z} \rightarrow \mathbf{Z}/n\mathbf{Z}$ son las aplicaciones canónicas.

f es un homomorfismo de anillos; el núcleo α de f está formado por los enteros x tales que $p(x) = 0$ y $q(x) = 0$, lo que significa: x es múltiplo común a m y n , y puesto que m y n son primos entre sí, su mcm es mn , luego $\alpha = mn\mathbf{Z}$.

Según III.4.4, la imagen de f es un anillo isomorfo a $\mathbf{Z}/mn\mathbf{Z}$; $\mathbf{Z}/mn\mathbf{Z}$ y $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ poseen el mismo cardinal mn , lo cual permite ver que f es epiyectiva. Finalmente, los anillos $\mathbf{Z}/mn\mathbf{Z}$ y $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ son isomorfos.

Designemos por $G(a)$ al grupo de las unidades de $\mathbf{Z}/a\mathbf{Z}$. Sabemos que $G(m) \times G(n)$ es el grupo de las unidades de $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$. Poniendo

$$\varphi(a) = \text{card}(G(a))$$

(función de Euler), de lo que precede se deduce la relación:

$$(4) \quad \varphi(mn) = \varphi(m) \varphi(n),$$

válida para m y n primos entre sí.

Cuando p es un entero primo y α es un entero ≥ 1 , $\varphi(p^\alpha)$ es el número de los enteros $< p^\alpha$ que no son múltiplos de p ; los números $\leq p^\alpha$ que son múltiplos de p son los mp , con $m \leq p^{\alpha-1}$; su número es, pues, $p^{\alpha-1}$, lo que da:

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^{\alpha-1}(p - 1).$$

Con la ayuda de (4), se deduce de (5) el siguiente resultado, debido a Euler:

III.4.6 Si la descomposición del entero n en factores primos distintos es:

$$\left\| \begin{array}{l} n = p_1^{\alpha_1} \dots p_k^{\alpha_k} \quad \alpha_1 \geq 1, \dots, \alpha_k \geq 1, \\ \text{se tiene:} \\ \varphi(n) = \prod_{i=1}^k p_i^{\alpha_i-1} (p_i - 1) \\ \text{en donde } \varphi(n) \text{ designa el número de enteros } < n \text{ y primos con } n. \end{array} \right.$$

§ III.5 DIVISIBILIDAD EN UN ANILLO

● En lo que sigue, consideraremos un anillo A conmutativo, unífero e íntegro.

DEFINICIÓN III.5.1

$\left\{ \begin{array}{l} \text{Sean } a, b \text{ dos elementos del anillo } A. \text{ Se dice que } a \text{ divide a } b \text{ (o} \\ \text{que } b \text{ es divisible por } a, \text{ o que } b \text{ es múltiplo de } a), \text{ si existe un } \lambda \in A \text{ tal} \\ \text{que } b = \lambda a. \\ \text{Entonces se escribe } a \mid b. \end{array} \right.$

Propiedades inmediatas

- Para todo $a \in A$, $a \mid a$ (pues $a = 1 \cdot a$).
- $a \mid b$ y $b \mid c$ implican $a \mid c$ (transitividad de la relación $a \mid b$).
- Todo $a \in A$ es múltiplo de 1, y todo $a \in A$ divide a 0.
- Sea $a \neq 0$, $a \in A$. Si $b \mid a$ y $a \mid b$, existen λ y $\mu \in A$ tales que $b = \lambda a$ y $a = \mu b$, de donde $a = \lambda \mu a$, y, puesto que A es íntegro, $\lambda \mu = 1$. Recíprocamente, para todo elemento invertible λ y todo $a \in A$, $(\lambda a) \mid a$ y $a \mid (\lambda a)$.

DEFINICIÓN III.5.2

$\left\{ \begin{array}{l} \text{Dos elementos } a \text{ y } b \text{ del anillo } A \text{ son } \textbf{asociados} \text{ si existe un elemento} \\ \text{invertible } \lambda \text{ tal que } b = \lambda a. \end{array} \right.$

El conjunto de elementos invertibles de A es un grupo respecto de la multiplicación (el grupo de las unidades), por lo que se tiene:

III.5.1 La relación « a y b son asociados» es una relación de equivalencia en \parallel el anillo A .

A las clases de equivalencia de esta relación se les llama *clases de los elementos asociados*.

La relación $a \mid b$ no es una relación de orden en A ; se obtiene una relación de orden si se pasa a *ideales*. Recordemos que, para todo $a \in A$, designamos por (a) al ideal principal engendrado por a : (a) está formado precisamente por los múltiplos de a (cf. p. 104).

III.5.2 \parallel Las relaciones « a y b son asociados» y « $(a) = (b)$ » son equivalentes.

Demostración. Si a y b están asociados, es inmediato que $(a) = (b)$.

Si $(a) = (b)$, se tiene: $a \mid b$ y $b \mid a$, luego a y b son asociados según el razonamiento que precede a la definición III.5.2. c.q.d.

Si hacemos corresponder a cada ideal principal el conjunto de sus generadores, se obtiene una *biyección canónica del conjunto de ideales principales de A , en el conjunto de clases de elementos asociados*.

DEFINICIÓN III.5.3

§ Se dice que el ideal principal (a) **divide** al ideal principal (b) si $(a) \mid (b)$.

Según las definiciones, si (a) divide a (b) , $a \mid b$, y recíprocamente. Puesto que la inclusión es una relación de orden en el conjunto $\mathcal{P}(A)$, tenemos:

TEOREMA III.5.3

|| Designemos por \mathcal{P} al conjunto de ideales principales del anillo A . En \mathcal{P} , la relación « a divide a b » es una relación de orden (opuesta a la relación de inclusión).

En todo lo que sigue dotaremos a \mathcal{P} de esta relación de orden.

DEFINICIÓN III.5.4

§ El elemento $a \neq 0$ del anillo A es **irreducible** si
 § 1.º $(a) \neq A$;
 § 2.º el ideal (a) es **minimal** en el conjunto $\mathcal{P} \setminus \{A\}$ respecto de la relación « a divide a b » (luego maximal respecto de la relación de inclusión).

Decir que a es irreducible es lo mismo que decir en primer lugar que a no es invertible, y seguidamente que los únicos divisores de a son los elementos invertibles de A por una parte, y por otra los elementos asociados a a .

DEFINICIÓN III.5.5

§ Sea $(a_i)_{i \in I}$ una familia de elementos del anillo A .
 § a) Se dice que la familia $(a_i)_{i \in I}$ admite un **máximo común divisor** en el anillo A , si la familia de ideales (a_i) admite un **ínfimo** δ en \mathcal{P} . Cuando esto ocurre, a todo generador de δ se le llama **un mcd** de los a_i ; notación: $\text{mcd}((a_i)_{i \in I})$.
 § b) Se dice que la familia $(a_i)_{i \in I}$ admite un **mínimo común múltiplo** en el anillo A , si la familia de ideales (a_i) admite un **supremo** \mathfrak{m} en \mathcal{P} .

§ Cuando esto ocurre, a todo generador de m se le llama **un mcm** de los a_i ; notación: $mcm((a_i)_{i \in I})$.

Según III.5.2, el mcd y el mcm están definidos a menos de un factor invertible.

Un anillo íntegro en el que todos sus ideales son principales se llama **anillo principal**. El teorema III.4.3 implica inmediatamente la siguiente propiedad:

TEOREMA III.5.4

|| En un anillo principal, toda familia de elementos admite un mcd y un mcm.

Existen anillos no principales en que esta propiedad también es verdadera (por ejemplo los anillos $K[X_1, \dots, X_n]$, cf. Cap. XIV).

Es preciso tener en cuenta que, en tales anillos, el ideal engendrado por el mcd de una familia de elementos $(a_i)_{i \in I}$ no coincide en general, con el ideal $\sum_{i \in I} (a_i)$, que es el ínfimo de los (a_i) en el conjunto de *todos los ideales* (ordenado por la relación de orden opuesta a la inclusión). Se puede hacer una observación análoga con el ideal engendrado por el mcm de los $(a_i)_{i \in I}$.

Así, en el anillo $A = K[X, Y]$ (en donde K designa a un cuerpo conmutativo), el mcd de los elementos X e Y es 1. Sin embargo, el ideal $(X) + (Y)$, formado por los polinomios sin término constante, es distinto de A .

Ejemplos

1) El anillo \mathbf{Z} es principal (cf. aplicación 1) de III.4.4; el grupo de las unidades de \mathbf{Z} es $\{-1, +1\}$. Las clases de elementos asociados de \mathbf{Z} son los conjuntos $\{-n, n\}$, $n \in \mathbf{N}$.

Los elementos irreducibles de \mathbf{Z} son los números $\pm p$, en donde $p \in \mathbf{N}^*$ es primo.

2) Si K es un cuerpo conmutativo, el anillo $K[X]$ de los polinomios en una variable, con coeficientes en K , es principal. Este anillo se estudiará detalladamente en el capítulo IV.

3) (Cf. Cap. IV) sea K un cuerpo conmutativo, y $K[X, Y]$ el anillo de los polinomios en dos variables sobre K .

$K[X, Y]$ no es principal; por ejemplo, el ideal (X, Y) no es principal, puesto que los polinomios que lo componen carecen de término constante, y que los únicos divisores comunes a X y a Y son las constantes $\neq 0$, que no pertenecen al ideal. Los elementos invertibles de $K[X, Y]$ son las constantes $\neq 0$.

Probemos que en $K[X, Y]$ el polinomio $P = X^2 + Y^2 + 1$ es, en general, irreducible. Si P no fuese irreducible, existirían constantes

$$u, v, r, u', v', r'$$

tales que

$$X^2 + Y^2 + 1 = (uX + vY + r)(u'X + v'Y + r'),$$

de donde

$$(1) \quad uu' = vv' = rr' = 1$$

$$(2) \quad ur' + u'r = vr' + v'r = 0$$

$$(3) \quad uv' + vu' = 0.$$

(1) prueba que u, v, r, u', v', r' son no nulos. (2) y (3) dan

$$\frac{u}{v} = -\frac{u'}{v'}, \quad \frac{v}{r} = -\frac{v'}{r'}, \quad \frac{r}{u} = -\frac{r'}{u'}$$

de donde, multiplicando miembro a miembro: $1 = -1$; luego K es de característica 2. Pero entonces, según (1), $uu' + vv' = 0$, de donde junto con (3):

$$(u + v)(u' + v') = 0, \quad \text{y} \quad u = -v = v, \quad \text{y además} \quad v = r, \quad u' = v' = r'.$$

Luego P es irreducible si K no es de característica 2; y si K es de característica 2, se tiene:

$$X^2 + Y^2 + 1 = (X + Y + 1)^2.$$

Se puede demostrar, sin embargo (cf. Cap. XIV) que todo polinomio de $K[X, Y]$ es producto de polinomios irreducibles, siendo la descomposición única salvo para los factores irreducibles. En general, este resultado es verdadero en el anillo

$$K[X_1, X_2, \dots, X_n] \quad (\text{cf. § IV.7}).$$

§ III.6 CUERPOS

DEFINICIÓN III.6.1

$\left\{ \begin{array}{l} \text{Un } \mathbf{cuerpo} \text{ es un anillo unífero en el cual todo elemento no nulo es} \\ \text{invertible. Si la multiplicación de un cuerpo es conmutativa, se dice que} \\ \text{el cuerpo es } \mathbf{conmutativo}. \end{array} \right.$

Se puede decir también que un cuerpo es un anillo unífero en que el grupo de unidades es el conjunto de elementos $\neq 0$.

En particular, un cuerpo es un anillo íntegro. Resulta de ello que *la característica de un cuerpo es 0, o un número primo* (cf. § 4).

DEFINICIÓN III.6.2

Sea K un cuerpo. Se llama **subcuerpo** de K a todo subanillo unífero de K , cuya estructura sea una estructura de cuerpo.

— Para comprobar que una parte L del cuerpo K es un subcuerpo de K , es suficiente ver:

- 1) que $-1 \in L$;
- 2) que L es estable para la suma y el producto de K ;
- 3) que la relación $x \in L$ implica $x^{-1} \in L$ si $x \neq 0$ (x^{-1} , inverso de x en K).

— Si L es un subcuerpo de K , se dice también que K es un *supercuerpo* de L .

— Si K y L son, además, conmutativos, se dice que L es una **extensión** de K .

TEOREMA III.6.1

Sea K un cuerpo, A un anillo unífero y $\rho : K \rightarrow A$ un homomorfismo de anillos. Si ρ es no nulo, ρ es inyectivo, y el anillo $\rho(K)$ es un cuerpo.

Demostración. Si $\rho \neq 0$, $\rho(1) = 1$. Sea $x \in K$ y $x' \in K$ tales que

$$\rho(x) = \rho(x').$$

Haciendo $u = x - x'$, se obtiene $\rho(u) = 0$. Supongamos que $u \neq 0$; se tiene entonces:

$$\rho(u \cdot u^{-1}) = \rho(1) = 1 = \rho(u) \rho(u^{-1}) = 0,$$

lo cual es absurdo. Luego $u = 0$ y $x = x'$, lo que demuestra que ρ es inyectivo. Es claro entonces que $\rho(K)$ es un cuerpo isomorfo a K . c.q.d.

TEOREMA III.6.2

|| *Todo anillo unífero, íntegro y finito es un cuerpo.*

Demostración. Sea A un anillo como el descrito, y sea $a \neq 0$, $a \in A$. Sean γ_a la aplicación $x \mapsto ax$ y δ_a la aplicación $x \mapsto xa$. γ_a y δ_a son inyecciones de A en A .

luego son biyecciones puesto que A es finito; luego existe un b tal que $ab = 1$ y un c tal que $ca = 1$, lo que demuestra que a es invertible. c.q.d

Nota. Se demuestra que tales cuerpos son necesariamente conmutativos (T. de Wedderburn: *todo cuerpo finito es conmutativo*) (cf. ejercicios: Cap. IV, polinomios ciclotómicos).

Aplicaciones

1) Si p es un entero > 0 y primo, el anillo $\mathbf{Z}/p\mathbf{Z}$ es un cuerpo.

En general, si A es un anillo unífero íntegro de característica $p > 0$, el subanillo engendrado por 1 en A es un cuerpo K isomorfo a $\mathbf{Z}/p\mathbf{Z}$: este cuerpo es el menor cuerpo contenido en A . Cuando A es un cuerpo, a K se le llama el *subcuerpo primo* de A .

2) Si K es un cuerpo de característica nula, sabemos ya que (en cuanto anillo) «contiene» a \mathbf{Z} ; contiene, pues, todos los elementos de la forma $(q.1)^{-1} \cdot (p.1)$, en donde $p \in \mathbf{Z}$ y $q \in \mathbf{Z}^*$. El conjunto de estos elementos forma un subcuerpo de K isomorfo a \mathbf{Q} . Podemos, pues, afirmar que *todo cuerpo de característica nula «contiene» a \mathbf{Q} , o que es un supercuerpo de \mathbf{Q} .*

La notación ideal permite, en el caso conmutativo, caracterizar los anillos que son cuerpos:

TEOREMA III.6.3

|| Sea A un anillo conmutativo unífero. Para que A sea un cuerpo, es necesario y suficiente que los únicos ideales de A sean $\{0\}$ y A .

Demostración. Si A es un cuerpo, sea \mathcal{I} un ideal $\neq \{0\}$ de A , y sea $a \in \mathcal{I}$, $a \neq 0$; puesto que \mathcal{I} es un ideal, se tiene: $1 = a^{-1} \cdot a \in \mathcal{I}$, luego $\mathcal{I} = A$.

Recíprocamente, sea A un anillo conmutativo en que $\{0\}$ y A son los únicos ideales, y sea $a \in A$, $a \neq 0$; el ideal (a) es no nulo, luego

$$(a) = A, \quad \text{y} \quad 1 \in (a),$$

lo que significa que a es invertible. c.q.d.

Como aplicación, vamos a caracterizar los **ideales maximales**, es decir, los **elementos maximales** del conjunto de los ideales de A , distintos de A , ordenado por inclusión. (A designa todavía un anillo conmutativo unífero.)

III.6.4 Si α designa un ideal del anillo A y $p: A \rightarrow A/\alpha$ la aplicación canónica, la aplicación $\mathfrak{b} \mapsto p(\mathfrak{b})$ es una biyección del conjunto de los ideales de A que contienen a α , en el conjunto de los ideales de A/α .

Demostración (resumida). Las relaciones « \mathfrak{b} es un ideal de A y $\mathfrak{b} \supset \mathfrak{a}$ » y « \mathfrak{b} es un ideal de A y $\mathfrak{b} = p^{-1}(p(\mathfrak{b}))$ » son equivalentes. \square

TEOREMA III.6.5

|| Para que el anillo cociente A/\mathfrak{a} sea un cuerpo, es necesario y suficiente que el ideal \mathfrak{a} sea **maximal** en el conjunto de los ideales de A **distintos de A , ordenado por inclusión.**

Demostración. Según III.6.3-4, A/\mathfrak{a} es un cuerpo si, y sólo si, los únicos ideales de A que contienen a \mathfrak{a} son A y \mathfrak{a} , o dicho de otra manera, si \mathfrak{a} es maximal en el conjunto de ideales $\neq A$, c.q.d.

Ejemplos de cuerpos

1) El conjunto \mathbf{Q} de los números racionales, dotado de la adición y de la multiplicación ordinarias, es un cuerpo. Asimismo, el conjunto \mathbf{R} de los números reales, y el conjunto \mathbf{C} de los números complejos; \mathbf{R} es un subcuerpo de \mathbf{C} , \mathbf{Q} es un subcuerpo de \mathbf{R} . Existe una infinidad de cuerpos K tales que

$$\mathbf{Q} \subset K \subset \mathbf{R}, \quad \text{o} \quad \mathbf{Q} \subset K \subset \mathbf{C}.$$

Por ejemplo, el conjunto K de los reales de la forma $a + b\sqrt{3}$, en donde $a \in \mathbf{Q}$ y $b \in \mathbf{Q}$, es un subcuerpo de \mathbf{R} que contiene a \mathbf{Q} : en efecto, puesto que $\sqrt{3} \notin \mathbf{Q}$, las relaciones $a \in \mathbf{Q}$, $b \in \mathbf{Q}$ y $a + b\sqrt{3} = 0$ implican $a = b = 0$. Es evidente que K es un subgrupo aditivo de \mathbf{R} , y que $1 \in K$ (hacer $a = 1$, $b = 0$). K es estable respecto del producto, ya que

$$(a + b\sqrt{3})(a' + b'\sqrt{3}) = aa' + 3bb' + (ab' + ba')\sqrt{3};$$

Finalmente sea $x \in K - \{0\}$, $x = a + b\sqrt{3}$; haciendo $\bar{x} = a - b\sqrt{3}$, se tiene

$$\bar{x} \neq 0 \quad (\text{puesto que } a \neq 0 \text{ y } b \neq 0), \quad \text{y} \quad x\bar{x} = a^2 - 3b^2 \neq 0,$$

de donde
$$\frac{1}{x} = \frac{a - b\sqrt{3}}{a^2 - 3b^2} \in K.$$

2) Sea p un entero > 0 y primo, y $\chi: \mathbf{Z} \rightarrow \mathbf{Z}/p\mathbf{Z}$ el homomorfismo canónico; $\mathbf{Z}/p\mathbf{Z}$ es un cuerpo con p elementos. El grupo de unidades $G(p)$ de este cuerpo tiene, pues, $p - 1$ elementos: $G(p) = \{\chi(1), \chi(2), \dots, \chi(p - 1)\}$. Sea

$$a \in G(p);$$

el orden de a en $G(p)$ divide a $p - 1$, lo que implica: $a^{p-1} = \chi(1)$. Esta propiedad es equivalente al «pequeño» teorema de Fermat: «para todo entero m no divisible por el número primo p , se tiene: $m^{p-1} \equiv 1(p)$ », quedando así demostrado. Se puede observar también que

$$(a^{p-1} = \chi(1)) \Rightarrow (a^p = a).$$

En consecuencia, la aplicación $x \mapsto x^p$ es la aplicación idéntica del cuerpo $\mathbf{Z}/p\mathbf{Z}$.

Ejemplos de ideales maximales

1) En todo anillo principal A , si q es un elemento irreducible, el ideal (q) es maximal (cf. § 5). Estos ideales son los únicos ideales maximales de A .

2) Sea K un cuerpo conmutativo, E un conjunto; designemos por A el anillo $\mathcal{F}(E, K)$ de las aplicaciones de E en K , y por x_0 un elemento cualquiera de E , fijo. La aplicación $\tilde{x}_0: A \mapsto K, f \mapsto f(x_0)$ es un homomorfismo de anillos.

\tilde{x}_0 es epiyectiva, pues para todo $y \in K$, existe un $f \in A$ tal que $f(x_0) = y$, por ejemplo la función f definida por $f(x_0) = y$ y $f(x) = 0$ si $x \neq x_0$. El núcleo de \tilde{x}_0 es el ideal \mathfrak{M}_{x_0} de los f tales que $f(x_0) = 0$; por lo tanto, \mathfrak{M}_{x_0} es un ideal maximal de A , y el cociente A/\mathfrak{M}_{x_0} es un cuerpo isomorfo a K .

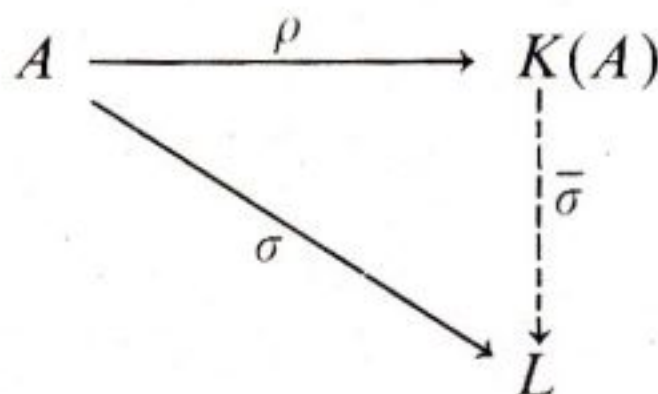
* Cuerpo de fracciones de un anillo unífero, conmutativo, íntegro

Vamos a generalizar el procedimiento de construcción de \mathbf{Q} a partir de \mathbf{Z} .

TEOREMA III.6.6

Sea A un anillo, unífero, conmutativo e íntegro. Existe un cuerpo $K(A)$ y un homomorfismo inyectivo $\rho: A \rightarrow K(A)$ que posee la propiedad siguiente: (llamada «universal»).

Para todo cuerpo L , y todo homomorfismo inyectivo $\sigma: A \rightarrow L$, existe un homomorfismo $\bar{\sigma}: K(A) \rightarrow L$ **único** tal que $\sigma = \bar{\sigma} \circ \rho$:



$K(A)$ es conmutativo.

Además, dos cuerpos que verifiquen estas condiciones son isomorfos.

Este teorema significa dos cosas:

- a) la existencia de un cuerpo $K(A)$ que contiene a A ;
- b) pero también que $K(A)$ es «el más pequeño posible» (propiedad universal) lo cual nos asegura su «unicidad salvo para un isomorfismo».

Demostración

1) *Unicidad.* Suponemos que los dos cuerpos $K(A)$, $K'(A)$ y las dos inyecciones $\rho : A \rightarrow K(A)$, $\rho' : A \rightarrow K'(A)$ responden a las condiciones impuestas. Por hipótesis, existen homomorfismos

$$\bar{\rho}' : K(A) \rightarrow K'(A) \quad \text{y} \quad \bar{\rho} : K'(A) \rightarrow K(A)$$

tales que $\rho = \bar{\rho} \circ \rho'$ y $\rho' = \bar{\rho}' \circ \rho$. De ellos se deducen las relaciones

$$\bar{\rho} \circ \bar{\rho}' \circ \rho = \bar{\rho} \circ \rho' = \rho \quad \text{y} \quad \bar{\rho}' \circ \bar{\rho} \circ \rho' = \bar{\rho}' \circ \rho = \rho'.$$

Por otro lado, si I e I' designan las aplicaciones idénticas de $K(A)$ y $K'(A)$

$$I \circ \rho = \rho = \bar{\rho} \circ \bar{\rho}' \circ \rho,$$

luego, en virtud de la propiedad universal, $\bar{\rho} \circ \bar{\rho}' = I$. Análogamente, $\bar{\rho}' \circ \bar{\rho} = I'$. Por lo tanto, $\bar{\rho}$ y $\bar{\rho}'$ son isomorfismos recíprocos.

2) Existencia

a) Construcción de ρ y de $K(A)$.

Sea \mathcal{K} el conjunto de los pares (a, b) , $a \in A$, $b \in B$ y $b \neq 0$. La relación binaria R definida en \mathcal{K} por

$$(a, b) R (a', b') \Leftrightarrow ab' = ba',$$

es una relación de equivalencia (su transitividad resulta del hecho de que A sea íntegro); designemos por $K(A)$ al conjunto cociente \mathcal{K}/R , y por $p : \mathcal{K} \rightarrow \mathcal{K}/R$ a la aplicación canónica. Dotamos a \mathcal{K} de las leyes siguientes (que están definidas puesto que A es íntegro):

$$\text{adición} \quad (a, b) + (c, d) = (ad + bc, bd),$$

$$\text{multiplicación} \quad (a, b) \cdot (c, d) = (ac, bd).$$

R es compatible con estas leyes. Por ejemplo, comprobémoslo para la adición; esta ley es conmutativa, luego es suficiente probar que si $(a, b) R (a', b')$, se tiene:

$$[(a, b) + (c, d)] R [(a', b') + (c, d)].$$

Esto nos lleva a comprobar la relación $(ad + bc) b' d = bd(a' d + b' c)$, o sea

$$ab' d^2 + bb' cd = a' bd^2 + bb' cd,$$

que es verdadera puesto que $ab' = ba'$.

Por paso al cociente, las leyes de \mathcal{K} definen leyes internas en $K(A)$. Según la nota que sigue a la definición II.2.4, estas leyes son asociativas y conmutativas (pues es fácil comprobar estas propiedades en \mathcal{K}).

$K(A)$ es un grupo abeliano respecto de la adición: el elemento neutro es $p((0, 1))$, el opuesto de $p((a, b))$ es $p((-a, b))$.

Para la multiplicación, $K(A)^* = K(A) \setminus \{0\}$ es un grupo abeliano: el elemento neutro es $p((1, 1))$, el inverso de (a, b) es (b, a) ($a \neq 0$).

La multiplicación es distributiva respecto de la adición (es suficiente comprobar esta propiedad en \mathcal{K}).

En resumen, $K(A)$ es un cuerpo conmutativo.

Si a y b son elementos de A , $b \neq 0$, el elemento $p((a, b))$ se designa por $\frac{a}{b}$. Se

define el homomorfismo $\rho: A \rightarrow K(A)$ por $\rho(a) = \frac{a}{1}$ ($a \in A$); es evidente que ρ es inyectiva.

b) Comprobación de la propiedad universal.

Sea L un cuerpo, y $\sigma: A \rightarrow L$ un homomorfismo inyectivo. Si existe una prolongación $\bar{\sigma}$ de σ a $K(A)$, debe cumplir:

$$\bar{\sigma}\left(\frac{a}{b}\right) = \bar{\sigma}\left(\frac{a}{1}\right) \bar{\sigma}\left(\frac{1}{b}\right) = \bar{\sigma}\left(\frac{a}{1}\right) \left[\bar{\sigma}\left(\frac{b}{1}\right)\right]^{-1} = \frac{\sigma(a)}{\sigma(b)},$$

lo que determina a $\bar{\sigma}$ de forma única. Recíprocamente, la aplicación

$$\bar{\sigma}: K(A) \rightarrow L$$

tal que $\bar{\sigma}\left(\frac{a}{b}\right) = \frac{\sigma(a)}{\sigma(b)}$ está definida, pues $\frac{\sigma(a)}{\sigma(b)}$ depende únicamente de la clase (a, b) .

En efecto:

$$(b \neq 0) \Rightarrow (\sigma(b) \neq 0), \quad \text{y} \quad \left(\frac{a}{b} = \frac{a'}{b'}\right) \Rightarrow (\sigma(a) \sigma(b') = \sigma(a') \sigma(b)),$$

de donde

$$\bar{\sigma}\left(\frac{a}{b}\right) = \bar{\sigma}\left(\frac{a'}{b'}\right).$$

Para terminar, se comprueba que $\bar{\sigma}$ es un homomorfismo. c.q.d.

DEFINICIÓN III.6.3

$\left\{ \begin{array}{l} \text{Se llama } \mathbf{cuerpo \ de \ fracciones} \text{ de un anillo íntegro conmutativo } A, \\ \text{al cuerpo } K(A) \text{ construido en la demostración del teorema III.6.4.} \\ \text{A la inyección } \rho: A \rightarrow K(A) \text{ se le llama inyección canónica.} \end{array} \right.$

Casi siempre, con la ayuda de ρ , se identifica A a un subanillo de $K(A)$. Se ve, pues, que *todo anillo íntegro conmutativo puede considerarse como un subanillo de un cuerpo*.

§ III.7 CÁLCULO EN EL CUERPO DE LOS NÚMEROS COMPLEJOS

Recordaremos brevemente la construcción y las propiedades esenciales del cuerpo \mathbf{C} de los números complejos.

En el conjunto $\mathbf{R} \times \mathbf{R}$ se obtiene una estructura de *cuerpo conmutativo* si se define la suma por medio de la fórmula: $(x, y) + (x', y') = (x + x', y + y')$ y la multiplicación por la fórmula $(x, y) \cdot (x', y') = (xx' - yy', xy' + x'y)$. El cuerpo así obtenido se designa por \mathbf{C} y se llama *cuerpo de los números complejos*.

El elemento nulo de \mathbf{C} es $(0, 0)$, el elemento unidad es $(1, 0)$; el inverso del elemento no nulo (x, y) es

$$\left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right).$$

La aplicación $j: \mathbf{R} \rightarrow \mathbf{C}$ tal que $j(x) = (x, 0)$ para todo $x \in \mathbf{R}$, es un isomorfismo del cuerpo \mathbf{R} en un subcuerpo de \mathbf{C} . Habitualmente, \mathbf{R} se identifica con el subcuerpo $j(\mathbf{R})$ de \mathbf{C} obtenido por medio de j , de suerte que el número complejo $(x, 0)$ se designa simplemente por x .

El elemento $(0, 1)$ de \mathbf{C} se designa por i ; se tiene: $i^2 = -1$. Por definición, el número complejo $z = (x, y)$ es igual a $x + iy$, y esta representación es única. Se establece:

$$x = \operatorname{Re}(z) \quad (\text{parte real de } z)$$

$$y = \operatorname{Im}(z) \quad (\text{parte imaginaria de } z).$$

Si $z = x + iy$ es un número complejo, se escribe $\bar{z} = x - iy$, y a \bar{z} se le llama el *conjugado* de z . Se comprueba inmediatamente la siguiente propiedad:

III.7.1 La aplicación $z \mapsto \bar{z}$ es un automorfismo involutivo de \mathbf{C} , que deja fijo
 \parallel cada uno de los elementos de \mathbf{R} .

Recíprocamente, un automorfismo de \mathbf{C} que deje fijo cada número real es la identidad si deja fijo a i ; si no, transforma a i en $-i$, y es la aplicación $z \mapsto \bar{z}$.

DEFINICIÓN III.7.1

El **módulo** del número complejo $z = x + iy$ es el número positivo

$$|z| = \sqrt{z \cdot \bar{z}} = \sqrt{x^2 + y^2}.$$

Se designa por $|z|$.

Se tiene evidentemente $z\bar{z} = |z|^2$, y, si $z \neq 0$, $z^{-1} = \frac{\bar{z}}{|z|^2}$.

Además

$$|z| \geq |\operatorname{Re}(z)|, \quad |z| \geq |\operatorname{Im}(z)| \quad z + \bar{z} = 2 \operatorname{Re}(z), \quad z - \bar{z} = 2i \operatorname{Im}(z).$$

z es real si, y sólo si, $z = \bar{z}$.

III.7.2 Para todo $z \in \mathbf{C}$ y $z' \in \mathbf{C}$, se tiene:

$$\| (1) |zz'| = |z| \cdot |z'|, \quad y \quad (2) \quad |z + z'| \leq |z| + |z'|.$$

Demostración

(1) Se tiene $|zz'|^2 = zz' \overline{zz'} = |z|^2 |z'|^2$ de donde $|zz'| = |z| \cdot |z'|$

(2) $|z + z'|^2 = (z + z')(\bar{z} + \bar{z}') = |z|^2 + |z'|^2 + z'\bar{z} + z\bar{z}'$.

luego:

$$|z + z'|^2 \leq |z|^2 + |z'|^2 + 2|zz'| = (|z| + |z'|)^2$$

y

$$|z + z'| \leq |z| + |z'|. \quad \text{c.q.d.}$$

III.7.3 Todo número complejo $\neq 0$ posee 2 raíces cuadradas opuestas.

Demostración. Hagamos $a = \alpha + i\beta$ y $z = x + iy$ ($a \neq 0$):

$$(3) \quad (z^2 = a) \Leftrightarrow \begin{cases} x^2 - y^2 = \alpha \\ 2xy = \beta \\ x^2 + y^2 = \sqrt{\alpha^2 + \beta^2} \end{cases} \Leftrightarrow \begin{cases} x^2 = \frac{1}{2}(\sqrt{\alpha^2 + \beta^2} + \alpha) \\ y^2 = \frac{1}{2}(\sqrt{\alpha^2 + \beta^2} - \alpha) \\ 2xy = \beta. \end{cases}$$

A una determinación de x sólo corresponde una determinación de y , puesto que el signo de xy es el de β . Se obtienen, pues, dos pares (x, y) , $(-x, -y)$ que son solución de (3). c.q.d.

III.7.4 *El conjunto de números complejos de módulo 1 es un subgrupo del grupo multiplicativo \mathbf{C}^* de los números complejos no nulos.*

Demostración. La aplicación $\psi : \mathbf{C}^* \rightarrow \mathbf{R}_+^*$, definida por $\psi(z) = |z|$, es un homomorfismo de grupos (para las leyes multiplicativas), pues

$$|zz'| = |z| \cdot |z'|;$$

y el conjunto de los números complejos de módulo 1 es precisamente el núcleo de ψ . c.q.d.

El grupo multiplicativo de los números complejos de módulo 1 se representa por U .

Homomorfismo exponencial, ángulos orientados

En el volumen II de la presente obra (Análisis), demostraremos el teorema siguiente:

TEOREMA III.7.5

- a) Existe un homomorfismo continuo **epiyectivo** φ del grupo aditivo \mathbf{R} en el grupo multiplicativo U , definido por $\varphi(t) = \sum_{n \geq 0} \frac{(it)^n}{n!}$, y un número $\pi > 0$ tal que el núcleo de φ es el subgrupo $2\pi\mathbf{Z}$ de \mathbf{R} ;
- b) todo homomorfismo continuo ψ del grupo aditivo \mathbf{R} en el grupo multiplicativo U es de la forma $x \mapsto \varphi(ax)$, en donde a es un número real que depende de ψ .

Las funciones circulares se definen haciendo $\varphi(\theta) = \cos \theta + i \sin \theta$ y se tiene: $\cos^2 \theta + \sin^2 \theta = 1$; las funciones $\cos \theta$ y $\sin \theta$ son periódicas, reales, analíticas, y se tiene

$$\frac{d}{d\theta} (\sin \theta) = \cos \theta, \quad \frac{d}{d\theta} (\cos \theta) = -\sin \theta.$$

DEFINICIÓN III.7.2

Al homomorfismo φ definido en III.7.5 se le llama **exponencial** y se le designa por $\varphi(\theta) = e^{i\theta}$.

Con estas notaciones, se tiene:

$$e^{i(\theta+\theta')} = e^{i\theta} \cdot e^{i\theta'}; \quad (e^{i\theta})^n = e^{in\theta} \quad (\text{fórmula de De Moivre}).$$

Con la ayuda de este teorema se puede construir la teoría completa de las funciones circulares (cf. Tomo II).

Argumento

DEFINICIÓN III.7.3

Para todo $z \in \mathbf{C}^* = \mathbf{C} \setminus \{0\}$, se llama **argumento** de z a cualquiera de los números reales θ tales que $e^{i\theta} = \frac{z}{|z|}$.

Puesto que $\frac{z}{|z|} \in U$ y que la exponencial es epiyectiva, todo elemento de \mathbf{C}^* tiene, por lo menos, un argumento; y por definición de núcleo, la igualdad $e^{i\theta} = e^{i\theta_0}$, que equivale a $e^{i(\theta-\theta_0)} = 1$, se verifica si, y sólo si, $\theta - \theta_0 \in 2\pi\mathbf{Z}$. Luego si se conoce uno de los argumentos de z , por ejemplo θ_0 , los restantes se deducen sumándole un múltiplo entero de 2π .

Con estas definiciones todo número complejo no nulo se puede escribir en la forma, llamada *trigonométrica*:

$$z = r e^{i\theta},$$

en donde el número positivo $r = |z|$ está unívocamente determinado, y el número real $\theta = \arg z$ «definido salvo para $2k\pi$ » (en donde $k \in \mathbf{Z}$).

DEFINICIÓN III.7.4

Para todo $z \in \mathbf{C}^*$ se llama **medida del ángulo orientado** (\vec{Ox}, \vec{Oz}) al conjunto de los números reales $\arg z$; notación (Ox, Oz) .

Esta medida no es una función numérica, sino un isomorfismo del grupo de las rotaciones de centro O en el grupo aditivo $\mathbf{R}/2\pi\mathbf{Z}$.

Aplicaciones de la fórmula de De Moivre

1) En primer lugar se deducen las *fórmulas de Euler*:

$$(4) \quad \boxed{\cos \theta = \frac{1}{2} (e^{i\theta} + e^{-i\theta}), \quad \operatorname{sen} \theta = \frac{1}{2i} (e^{i\theta} - e^{-i\theta})},$$

luego

$$(e^{i\pi})^2 = e^{2i\pi} = 1, \text{ de donde } e^{i\pi} = \pm 1.$$

Puesto que $\pi \notin 2\pi\mathbf{Z}$, necesariamente $e^{i\pi} = -1$; el estudio de las funciones circulares prueba que $e^{i\frac{\pi}{2}} = i$, $e^{-i\frac{\pi}{2}} = -i$.

2) Suma de una progresión geométrica.

Hacemos

$$S = 1 + e^{i\theta} + \dots + e^{in\theta} = \sum_{k=0}^n e^{ik\theta} = \sigma + i\tau.$$

Si $e^{i\theta} = 1$ (e.d. si $\theta = 2k\pi$, $k \in \mathbf{Z}$), se tiene: $S = n + 1$.

Si $e^{i\theta} \neq 1$ (o sea $\theta \neq 2k\pi$), se tiene:

$$S = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}},$$

de donde

$$S = \frac{e^{i\frac{n+1}{2}\theta} (e^{i\frac{n+1}{2}\theta} - e^{-i\frac{n+1}{2}\theta})}{e^{i\frac{\theta}{2}} (e^{i\frac{\theta}{2}} - e^{-i\frac{\theta}{2}})} = e^{in\frac{\theta}{2}} \cdot \frac{\operatorname{sen} \frac{n+1}{2} \theta}{\operatorname{sen} \theta/2}.$$

Tomando las partes reales e imaginarias, se obtiene

(5)

con

$$\begin{aligned} \sigma &= \cos \frac{n\theta}{2} \frac{\operatorname{sen} \frac{n+1}{2} \theta}{\operatorname{sen} \theta/2}, & \tau &= \operatorname{sen} \frac{n\theta}{2} \frac{\operatorname{sen} \frac{n+1}{2} \theta}{\operatorname{sen} \theta/2} \\ \sigma &= \sum_{k=0}^n \cos k\theta; & \tau &= \sum_{k=0}^n \operatorname{sen} k\theta \end{aligned}$$

3) Linealización de $\cos^n x$.

Haremos el cálculo solamente para el caso $n = 2p$, p entero.

Según la fórmula de Euler, $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$, de donde

$$\cos^{2p} x = \frac{1}{2^{2p}} (e^{ix} + e^{-ix})^{2p}$$

Aplicando la fórmula del binomio y agrupando los términos equidistantes de los extremos, se obtiene

$$\cos^{2p} x = \frac{1}{2^{2p}} \left[\binom{2p}{p} + \sum_{k=0}^{p-1} \binom{2p}{k} (e^{2(k-p)ix} + e^{2(p-k)ix}) \right].$$

La aplicación de las fórmulas de Euler a cada uno de los términos del segundo miembro nos da finalmente

$$(6) \quad \cos^{2p} x = \frac{1}{2^{2p}} \left[\binom{2p}{p} + 2 \sum_{k=0}^{p-1} \binom{2p}{k} \cos 2(p-k)x \right].$$

De forma análoga se calcula $\cos^{2p+1} x$ y $\sin^n x$.

4) Expresión de $\cos nx$ como polinomio en $\cos x$.

Efectuamos el cálculo para $n = 2p$ ($p \in \mathbb{N}$).

Según la fórmula de De Moivre,

$$(7) \quad (\cos x + i \sin x)^{2p} = \cos 2px + i \sin 2px.$$

Aplicando la fórmula del binomio al primer miembro de (7), y tomando las partes reales de ambos miembros, se obtiene:

$$(8) \quad \begin{aligned} (\cos x + i \sin x)^{2p} &= \sum_{k=0}^{2p} \binom{2p}{k} \cos^k x (i \sin x)^{2p-k} \\ \cos 2px &= \sum_{m=0}^p \binom{2p}{2m} (-1)^{p-m} \cos^{2m} x \sin^{2(p-m)} x. \end{aligned}$$

Si cambiamos m por $p-m$ en (8), y teniendo en cuenta

$$(9) \quad \sin^{2m} x = (1 - \cos^2 x)^m = \sum_{\lambda=0}^m (-1)^\lambda \binom{m}{\lambda} \cos^{2\lambda} x,$$

se tiene:

$$\cos 2px = \sum_{0 \leq m \leq p, 0 \leq \lambda \leq m} (-1)^{m+\lambda} \binom{2p}{2m} \binom{m}{\lambda} \cos^{2(p-m+\lambda)} x,$$

que, cambiando λ por $m-\lambda$, es:

$$(10) \quad \cos 2px = \sum_{0 \leq \mu \leq p} (-1)^\mu \left(\sum_{\mu \leq m \leq p} \binom{2p}{2m} \binom{m}{\mu} \right) \cos^{2(p-\mu)} x.$$

Si hacemos $\mu = 0$ en esta última expresión, se encuentra el coeficiente de $\cos^{2p} x$, o sea

$$\sum_{0 \leq m \leq p} \binom{2p}{2m} = 2^{2p-1}.$$

En el capítulo IV, § 5, obtendremos una expresión más simple del segundo miembro de (10).

También es posible calcular la expresión $\frac{\operatorname{sen} nx}{\operatorname{sen} x}$ en función de $\cos x$.

Raíces de la unidad

III.7.6 *Todo número complejo no nulo posee n raíces n -ésimas distintas.*

Demostración. Se escribe $a = \rho e^{i\theta}$, $\alpha = r e^{i\varphi}$ ($\rho \geq 0$, $r \geq 0$). La ecuación: $\alpha^n = a$ equivale a $(r^n e^{in\varphi} = \rho e^{i\theta})$, o sea a

$$\begin{cases} r^n = \rho \\ n\varphi = \theta + 2k\pi, k \in \mathbf{Z}, \end{cases} \quad \text{o también a} \quad \begin{cases} r = \sqrt[n]{\rho}, & \varphi = \frac{\theta}{n} + \frac{2k\pi}{n} + 2\pi\lambda \\ k, \lambda \in \mathbf{Z}, & 0 \leq k \leq n-1 \text{ c.q.d.} \end{cases}$$

En particular, la ecuación $\zeta^n = 1$ ($n \in \mathbf{N}$) tiene n raíces distintas ζ_k dadas por $\zeta_k = e^{2i\frac{k\pi}{n}}$, $0 \leq k \leq n-1$. Estos n números son las *raíces n -ésimas de la unidad*. Observemos que toda raíz de la unidad ζ es un elemento de U , por lo tanto $\zeta\bar{\zeta} = 1$. Designaremos por U_n al conjunto de las raíces n -ésimas de 1 ⁽¹⁾.

III.7.7 U_n es un subgrupo del grupo multiplicativo \mathbf{C}^* , isomorfo al grupo aditivo $\mathbf{Z}/n\mathbf{Z}$.

Demostración. Hagamos $\omega = \zeta_1 = e^{2i\frac{\pi}{n}}$. La aplicación $\psi: m \rightarrow \omega^m$ de \mathbf{Z} en U_n es un homomorfismo *epiyectivo* de grupos, cuyo núcleo es el conjunto de los $m \in \mathbf{Z}$ tales que $\omega^m = 1$. Pero $\omega^m = 1$ se escribe $e^{2im\frac{\pi}{n}} = 1$, lo que equivale a la condición $m/n \in \mathbf{Z}$. Luego este núcleo es $n\mathbf{Z}$, y U_n es isomorfo a $\mathbf{Z}/n\mathbf{Z}$ según II.5.3. c.q.d.

Los generadores de $\mathbf{Z}/n\mathbf{Z}$ están representados por los enteros k que verifican

$$1 \leq k \leq n-1 \quad \text{y} \quad \operatorname{mcd}(k, n) = 1.$$

⁽¹⁾ Los especialistas en teoría de *grupos algebraicos* suelen designar a este conjunto por μ_n . Por razones de coherencia en la notación, en esta obra lo hemos llamado U_n .

Luego los generadores de U_n son los números ω^k tales que:

$$1 \leq k \leq n-1 \quad \text{y} \quad \text{mcd}(k, n) = 1.$$

DEFINICIÓN III.7.5

A los generadores de U_n se les llama raíces n -ésimas primitivas de la unidad.

En el capítulo IV, § 5, demostraremos el resultado siguiente (e incluso un resultado más general).

TEOREMA III.7.8

El único subgrupo finito de orden n del grupo U es U_n .

§ III.8 ⁽¹⁾ ESTRUCTURA DE MÓDULO SOBRE UN ANILLO

● Todos los anillos que se consideran en este párrafo son uníferos.

DEFINICIÓN III.8.1

Sea A un anillo; un A -módulo por la izquierda es un conjunto M provisto de una ley interna (designada aditivamente) y de una ley externa por la izquierda, de dominio A , designada por $(\alpha, x) \mapsto \alpha x$ ($\alpha \in A$, $x \in M$), tales que:

- (M₁) M es un grupo abeliano para la adición;*
- (M₂) la ley externa es distributiva respecto de la adición de A y respecto de la de M ;*
- (M₃) para todo $\alpha \in A$, todo $\beta \in A$ y todo $x \in M$, se tiene:*

$$\alpha(\beta x) = (\alpha\beta)x;$$
- (M₄) para todo $x \in M$, se tiene: $1.x = x$.*

Se puede definir la noción de A -módulo por la derecha (la ley externa es por la derecha) pero en esta obra no la necesitaremos.

Propiedades inmediatas

— Para todo $x \in M$, se tiene:

$$(1) \quad 0.x = 0,$$

⁽¹⁾ Este párrafo puede omitirse en una primera lectura.

pues $0.x = (0 + 0).x = 0.x + 0.x$, de donde $0.x = 0$ (atención: en la relación (1), el 0 del miembro de la izquierda es el de A , el 0 del miembro de la derecha es el de M).

— Para todo $\alpha \in A$ y todo $x \in M$,

$$(2) \quad (-\alpha).x = -(\alpha x)$$

pues $\alpha.x + (-\alpha).x = [\alpha + (-\alpha)].x = 0.x = 0$; en particular,

$$(-1).x = -(1.x) = -x.$$

— Para todo $\alpha \in A$, la aplicación $x \mapsto \alpha x$ es un endomorfismo del grupo aditivo M (según (M_2)). De lo que resulta $\alpha.0 = 0$. c.q.d.

Aplicaciones lineales

De acuerdo con las definiciones generales del capítulo II, se establece:

DEFINICIÓN III.8.2

Sean M, N dos A -módulos. Una aplicación $f: M \rightarrow N$ es un **homomorfismo** de A -módulos si es un homomorfismo para los grupos aditivos M y N , y si, para todo $\alpha \in A$ y todo $x \in M$, se tiene:

$$f(\alpha x) = \alpha f(x).$$

A los homomorfismos de A -módulos se les llama *aplicaciones A -lineales*, o *aplicaciones lineales* si sólo entra en juego un anillo de base.

La compuesta de dos aplicaciones lineales es una aplicación lineal.

El conjunto de las aplicaciones A -lineales de M en N se designa por $\mathcal{L}_A(M, N)$, o $\mathcal{L}(M, N)$ cuando no hay peligro de confusión. Se escribe $\mathcal{L}_A(M)$ en vez de $\mathcal{L}_A(M, M)$; $\mathcal{L}_A(M, N)$ es un grupo abeliano para la ley $(f, g) \mapsto f + g$ definida por:

$$(f + g)(x) = f(x) + g(x) \quad (x \in M).$$

Para $f \in \mathcal{L}_A(M, N)$ y $\alpha \in A$, sea $\alpha.f$ la aplicación de M en N tal que $(\alpha f)(x) = \alpha.f(x)$ para todo $x \in M$.

En general, αf no es lineal, puesto que

$$(\alpha f)(\lambda x) = \alpha.f(\lambda x) = \alpha(\lambda f(x)) = \alpha\lambda f(x),$$

y

$$\lambda(\alpha f)(x) = \lambda\alpha f(x).$$

Sin embargo, si A es conmutativo, αf es A -lineal, y la aplicación

$$(\alpha, f) \mapsto \alpha f$$

es una ley externa en $\mathcal{L}_A(M, N)$, de dominio A , que dota a $\mathcal{L}_A(M, N)$ de estructura de A -módulo.

Submódulos

— Sea M un A -módulo. Un *submódulo* de M es una parte de M estable para las leyes de M , y que es un A -módulo para las leyes inducidas.

TEOREMA III.8.1

|| Sea M un A -módulo y N una parte no vacía de M . Para que N sea un sub- A -módulo de M , es necesario y suficiente que N sea estable para la ley interna y para la ley externa de M .

Demostración. La condición es evidentemente necesaria. Recíprocamente, supongámosla verificada, y es suficiente probar que N es un subgrupo de M . Sea, pues, $x \in N$. Se tiene: $(-1).x \in N$; pero hemos visto antes que

$$(-1).x = -x,$$

de donde la relación: $(\forall x) x \in N \Rightarrow -x \in N$, lo que, en virtud de las hipótesis, prueba que N es un subgrupo de M . c.q.d.

Propiedades de los submódulos

— $\{0\}$ y M son submódulos de M . Si N es un submódulo, la inyección canónica $N \rightarrow M$ es lineal.

— Si $(N_i)_{i \in I}$ es una familia de sub- A -módulos de M , $N = \bigcap_{i \in I} N_i$ es un submódulo de M ; N es el ínfimo de los N_i en el conjunto, ordenado por inclusión, de los submódulos de M .

— En particular, la intersección de los submódulos que contienen a $\bigcup_{i \in I} N_i$ es un submódulo de M , supremo de los N_i ; este módulo se designa por $\sum_{i \in I} N_i$ y se llama *suma* de los N_i . Evidentemente, es el conjunto de elementos de M de la forma $\sum_{i \in J} x_i$, en donde J es una parte arbitraria, finita, no vacía, de I .

— Si M, N son dos A -módulos y $f \in \mathcal{L}_A(M, N)$, la imagen directa de todo submódulo de M por f es un submódulo de N , y la imagen recíproca de todo submódulo de N por f es un submódulo de M .

En particular, el núcleo $f^{-1}(0)$ de f es un submódulo de M , y la imagen $f(M)$ de f es un submódulo de N .

Combinaciones lineales en un módulo

Sea $(a_i)_{i \in I}$ una familia de elementos del A -módulo M . Recordemos (cf. Cap. I) que $i \mapsto a_i$ es una aplicación cualquiera (no necesariamente inyectiva) de I en M . El *conjunto asociado* es la parte de M imagen de I por medio de esta aplicación.

Recíprocamente, si S es una parte de M , es posible asociarle la familia $(u_s)_{s \in S}$ tal que, para todo $s \in S$, $u_s = s$. Se dice que esta familia está *canónicamente asociada* a S , o simplemente, *asociada a S* .

DEFINICIÓN III.8.3

Si $(a_i)_{i \in I}$ es una familia de elementos del A -módulo M , una **combinación lineal de los a_i** es un elemento de M de la forma:

$$(1) \quad x = \sum_{i \in I} \lambda_i a_i,$$

en donde $i \mapsto \lambda_i$ es una aplicación de I en A , tal que $\lambda_i = 0$ salvo para un conjunto finito de valores de i (brevemente, se dice que los λ_i son casi todos nulos).

A los λ_i se les llama **coeficientes** de la combinación lineal.

El sentido que debemos dar a (1) es el siguiente: si todos los λ_i son nulos, $x = 0$; en caso contrario, x es la suma $\sum_{i \in J} \lambda_i a_i$, en donde J designa a la parte (finita) de los $i \in I$ tales que $\lambda_i \neq 0$. Es cómodo convenir que el símbolo $\sum_{i \in \emptyset} \lambda_i a_i$ representa al elemento nulo de M . Con este convenio, en (1), x es la suma $\sum_{i \in J} \lambda_i a_i$, en donde J designa al conjunto de los $i \in I$ tales que $\lambda_i \neq 0$, en todos los casos.

Caso particular. Si S designa a una parte del A -módulo M , una combinación lineal de elementos de S es una combinación lineal de la familia asociada a S , es decir, un elemento $x \in M$ de la forma:

$$x = \sum_{s \in S} \lambda_s \cdot s,$$

en donde los coeficientes $(\lambda_s)_{s \in S}$ son casi todos nulos. c.q.d.

Consideremos una familia $(a_i)_{i \in I}$ de elementos del A -módulo M . El conjunto de submódulos N de M tales que $(\forall i \in I, a_i \in N)$ es no vacío, puesto que M verifica esta propiedad. La intersección de este conjunto de submódulos es *el menor submódulo de M que contiene a todos los a_i* , y se le llama **submódulo engendrado por la familia $(a_i)_{i \in I}$** .

III.8.2 El submódulo engendrado por una familia $(a_i)_{i \in I}$ de elementos del A -módulo M es el conjunto de las combinaciones lineales de los a_i .

Demostración. Según III.8.1, si un submódulo N de M contiene a todos los a_i , contiene toda combinación lineal de los a_i . Basta, pues, con demostrar que el conjunto \mathcal{N} de estas combinaciones lineales es un submódulo de M . Vamos a comprobar sucesivamente que si $x \in \mathcal{N}$ e $y \in \mathcal{N}$, se tiene: $x + y \in \mathcal{N}$; y que si $\lambda \in A$ y $x \in \mathcal{N}$, se tiene: $\lambda x \in \mathcal{N}$.

a) Si $x \in \mathcal{N}$ e $y \in \mathcal{N}$, existen escalares $(\lambda_i)_{i \in I}$ casi todos nulos, y escalares $(\mu_i)_{i \in I}$ casi todos nulos, tales que

$$x = \sum_{i \in I} \lambda_i a_i, \quad y = \sum_{i \in I} \mu_i a_i.$$

Los escalares $(\lambda_i + \mu_i)_{i \in I}$ son casi todos nulos, ya que si K (resp. L) designa una parte finita de I tal que $\lambda_i = 0$ para $i \notin K$ [resp. $\mu_i = 0$ para $i \notin L$], se tiene: $\lambda_i + \mu_i = 0$ para $i \notin K \cup L$, y $K \cup L$ es una parte finita de I . Es evidente que $x + y = \sum_{i \in I} (\lambda_i + \mu_i) a_i$, de donde $x + y \in \mathcal{N}$.

b) Si $x = \sum_{i \in I} \lambda_i a_i \in \mathcal{N}$, y si $\lambda \in A$, los escalares $(\lambda \lambda_i)_{i \in I}$ son casi todos nulos, y se tiene: $\lambda x = \sum_{i \in I} (\lambda \lambda_i) a_i$. c.q.d.

Si S designa una *parte* del A -módulo M , el submódulo engendrado por la familia asociada a S es el conjunto de las combinaciones lineales de elementos de S ; brevemente, se llamará **submódulo engendrado por S** . Este submódulo es, pues, el conjunto de elementos $x \in M$ de la forma:

$$x = \sum_{a \in S} \lambda_a \cdot a,$$

en donde $(\lambda_a)_{a \in S}$ es una familia de escalares casi todos nulos.

DEFINICIÓN III.8.4

Se dice que una parte S de un A -módulo M es una **parte generadora** (o un **sistema generador** o también, por abuso de lenguaje, un **sistema de generadores**) si el submódulo engendrado por S en M es igual a M .

III.8.3 Sean $(a_i)_{i \in I}$ y $(b_j)_{j \in J}$ dos familias de elementos del A -módulo M . Si el conjunto asociado a la familia $(a_i)_{i \in I}$, está **contenido** en el conjunto asociado a la familia $(b_j)_{j \in J}$, el submódulo N engendrado por la familia $(a_i)_{i \in I}$ está **contenido** en el submódulo P engendrado por la familia $(b_j)_{j \in J}$.

Demostración. Si $x \in N$, existen escalares $(\lambda_i)_{i \in I}$, casi todos nulos, tales que $x = \sum_{i \in I} \lambda_i a_i$. Sea K la parte finita de I formada por los $i \in I$ tales que $\lambda_i \neq 0$. Para todo $i \in K$, existe (según las hipótesis), por lo menos, un elemento de J , llamado $j(i)$ tal que $a_i = b_{j(i)}$; y se tiene:

$$x = \sum_{i \in K} \lambda_i a_i = \sum_{i \in K} \lambda_i b_{j(i)} \text{ de donde } x \in P. \text{ c.q.d.}$$

COROLARIO

|| Sea $(a_i)_{i \in I}$ una familia de elementos del A -módulo M . Toda familia $(b_j)_{j \in J}$ que tenga el mismo conjunto asociado engendra el mismo submódulo.

Demostración. Se aplica II.8.3 intercambiando los papeles de ambas familias.

DEFINICIÓN III.8.5

Sea $(a_i)_{i \in I}$ una familia de elementos del A -módulo M . Se dice que esta familia es **libre**, o que las a_i son **linealmente independientes**, si, cualquiera que sea la familia de escalares casi todos nulos, $(\lambda_i)_{i \in I}$, la relación:

$$\sum_{i \in I} \lambda_i a_i = 0$$

implica: $(\forall i \in I) \lambda_i = 0$.

Una parte L de M es **libre**, si la familia asociada a L es libre.

A una familia $(a_i)_{i \in I}$ de elementos de M , no libre, se le llama **ligada**, o **linealmente dependiente**.

III.8.4 Si la familia $(a_i)_{i \in I}$ de elementos del A -módulo M es libre, la aplicación $i \mapsto a_i$ de I en M es **inyectiva**. En otras palabras, cada elemento del conjunto asociado sólo figura una vez en la familia.

Demostración. Si $i \mapsto a_i$ no es una aplicación inyectiva, existen $i_0 \in I$ e $i_1 \in I$ tales que $i_0 \neq i_1$ y $a_{i_0} = a_{i_1}$. Hagamos $\lambda_{i_0} = -\lambda_{i_1} = 1$, y, para $i \neq i_0$ y $i \neq i_1$ ($i \in I$), $\lambda_i = 0$. Se tiene entonces $\sum_{i \in I} \lambda_i a_i = 0$, en donde los escalares (λ_i) son casi todos nulos, pero no todos nulos. Luego la familia $(a_i)_{i \in I}$ está ligada. c.q.d.

Llamamos *subfamilia* [resp. *subfamilia finita*] de una familia $(a_i)_{i \in I}$, a toda familia de la forma $(a_i)_{i \in J}$, en donde J es una parte [resp. una parte finita] de I . De las definiciones, resulta que:

III.8.5 Una familia $(a_i)_{i \in I}$ de elementos del A -módulo M es libre si, y sólo si,
 \parallel toda subfamilia finita de esta familia es libre.

Nota. $\{0\}$ no es una parte libre de M , puesto que, para todo $\lambda \in A$, $\lambda \cdot 0 = 0$. Luego, si $(a_i)_{i \in I}$ es una parte libre de M , para todo $i \in I$, se tiene: $a_i \neq 0$.

Sea $(a_i)_{i \in I}$ una familia libre del A -módulo M , y sea $(\lambda_i)_{i \in I}$ una familia de escalares casi todos nulos. Si $(\mu_i)_{i \in I}$ es otra familia de escalares casi todos nulos, tal que

$$\sum_{i \in I} \lambda_i a_i = \sum_{i \in I} \mu_i a_i,$$

se tiene

$$\sum_{i \in I} (\lambda_i - \mu_i) a_i = 0,$$

de donde $(\forall i \in I, \lambda_i = \mu_i)$, según la definición III.8.4.

Luego, si x es una combinación lineal de los a_i , existe una familia y sólo una de escalares $(\lambda_i)_{i \in I}$, casi todos nulos, tal que

$$x = \sum_{i \in I} \lambda_i a_i.$$

Al elemento λ_i se le llama entonces *coordenada de x de índice i* . Recíprocamente, si esta condición se verifica para toda combinación lineal de los a_i , entonces la familia $(a_i)_{i \in I}$ es libre.

DEFINICIÓN III.8.6

$\{$ Una familia $(a_i)_{i \in I}$ de elementos del A -módulo M es una **base** de M ,
 $\}$ si es, a la vez, libre y generadora.

Según lo que antecede, $(a_i)_{i \in I}$ es una base de M si, y sólo si, todo $x \in M$ se escribe de una manera y sólo una en la forma

$$x = \sum_{i \in I} \lambda_i a_i,$$

en donde los escalares (λ_i) son casi todos nulos.

Cuando esto ocurre, la aplicación $x \mapsto \lambda_i$ de M en A que es A -lineal se llama *proyección de índice i* . (Aquí A se halla dotado de su estructura canónica de A -módulo por la izquierda (cf. ejemplo 2, más abajo.))

Nota. Es cómodo dotar al conjunto de las partes generadoras (resp. de las partes de M , decir que $S \subset T$ equivale a decir que la familia asociada a S es una por la *inclusión* de las partes de M . Desde este punto de vista, si S y T designan a partes de M , decir que $S \subset T$ equivale a decir que la familia asociada a S es una subfamilia de la familia asociada a T . A lo largo de toda esta obra, nos permitiremos utilizar, sin mayores explicaciones, frases del tipo: «sean S y T bases de E tales que $S \subset T$ ».

Ejemplos de módulos

- 1) Sea G un grupo abeliano; la ley externa

$$(m, x) \mapsto m \cdot x \quad (m \in \mathbf{Z}, x \in G)$$

define en G una estructura de \mathbf{Z} -módulo. Los sub- \mathbf{Z} -módulos de G son los subgrupos de G . Todo teorema sobre los subgrupos abelianos es un teorema de los sub- \mathbf{Z} -módulos y viceversa.

- 2) Sea A un anillo; la ley «externa»

$$(\lambda, x) \mapsto \lambda \cdot x \quad (\lambda \in A, x \in A)$$

define en A una estructura (llamada *canónica*) de A -módulo. A los sub- A -módulos de A se les llama *ideales por la izquierda* de A . (Si A es conmutativo, éstos son los ideales ya definidos.)

- 3) Sea A un anillo y E un conjunto, y sea $\mathcal{F}(E, A)$ el anillo de las aplicaciones de E en A . $\mathcal{F}(E, A)$ es un A -módulo, si se le dota de la ley $(\alpha, f) \mapsto \alpha f$ tal que

$$(\alpha f)(x) = \alpha \cdot f(x) \quad (\alpha \in A, x \in E).$$

- 4) Sea G un grupo abeliano, E un conjunto, $\mathcal{F}(E, G)$ el grupo aditivo de las aplicaciones de E en G , y $\mathcal{E}(G)$ el anillo de los endomorfismos de G . La ley

$$(\alpha, f) \mapsto \alpha \circ f \quad (\alpha \in \mathcal{E}(G), f \in \mathcal{F}(E, G))$$

define en $\mathcal{F}(E, G)$ una estructura de $\mathcal{E}(G)$ -módulo por la izquierda.

Ejemplos de bases de A-módulos

1) Sea A un anillo; consideremos los n elementos e_i ($1 \leq i \leq n$) del módulo A^n , definidos por:

$$e_i = (\delta_{i,1}, \delta_{i,2}, \dots, \delta_{i,n}); \quad \delta_{i,i} = 1 \quad \text{y} \quad \delta_{i,j} = 0 \quad \text{si} \quad i \neq j \quad (1 \leq i \leq n);$$

$\{e_i\}_{1 \leq i \leq n}$ es una base de A^n , llamada *base canónica*.

2) Sea A un anillo, y sea $A^{(\mathbf{N})}$ el conjunto de las sucesiones $(x_n)_{n \geq 0}$ de elementos de A , nulas a partir de un cierto lugar; $A^{(\mathbf{N})}$ es un sub- A -módulo del módulo de las aplicaciones de \mathbf{N} en A . Establecemos

$$e_i = (\lambda_{i,n})_{n \geq 0}, \quad \lambda_{i,i} = 1 \quad \text{y} \quad \lambda_{i,j} = 0 \quad \text{para} \quad i \neq j \quad (i \in \mathbf{N});$$

los e_i forman una base (infinita) de $A^{(\mathbf{N})}$, llamada *base canónica*.

§ III.9 ESTRUCTURA DE ÁLGEBRA SOBRE UN ANILLO CONMUTATIVO UNÍFERO

En este párrafo, de forma excepcional, designaremos por $*$ a la multiplicación en ciertos anillos.

DEFINICIÓN III.9.1

Sea C un anillo conmutativo. Una C -álgebra asociativa es un C -módulo A , dotado de una multiplicación $*$ que hace de A un anillo unífero (cuyo elemento unidad designaremos por I), tal que, para todo $\lambda \in C$, todo $x \in A$, y todo $y \in A$, se tiene:

$$\lambda \cdot (x * y) = (\lambda x) * y = x * (\lambda y).$$

Nota. En esta obra, sólo consideraremos álgebras asociativas. Por este motivo, nos permitiremos llamarlas «álgebras» sin ningún calificativo.

Designamos por A una C -álgebra, por $*$ el producto de A .

— La aplicación $\psi : \lambda \mapsto \lambda \cdot I$ es un homomorfismo de anillos: en efecto $1 \cdot I = I$, y

$$(\lambda \cdot \mu) \cdot I = (\lambda \mu) (I * I) = \lambda \cdot (\mu (I * I)) = \lambda \cdot (I * (\mu I)) = (\lambda I) * (\mu I).$$

La imagen de C es un subanillo conmutativo de A , contenido en el *centro* de A ⁽¹⁾.

(1) Por definición, el centro de un anillo A es el conjunto de elementos $x \in A$ tales que $xy = yx$ para todo $y \in A$. Se comprueba fácilmente que el centro de A es un subanillo unífero conmutativo de A .

— Sea C un anillo unífero conmutativo, y sea D un subanillo unífero de C . Si A es una C -álgebra, el módulo obtenido por restricción a D de los escalares es una D -álgebra.

— Sea A un anillo unífero; el homomorfismo $m \mapsto m.I$ de \mathbf{Z} en A define en A una estructura de \mathbf{Z} -álgebra, llamada canónica.

— A una parte B de la subálgebra A se le llama *sub- C -álgebra de A* si es a la vez un sub- C -módulo de A y un subanillo unífero de A . Estas dos estructuras proveen, entonces, a B de una estructura de C -álgebra, llamada *inducida*.

— Sea A un anillo unífero, y C un subanillo de A contenido en el centro de A . La inyección canónica $j : C \rightarrow A$ define en A una estructura de C -álgebra.

Los ejemplos más importantes de álgebras son las álgebras de polinomios, de series formales, de matrices (Caps. IV, VII y VIII), y las álgebras de funciones, e.d. las subálgebras de $\mathcal{F}(E, A)$, en donde A designa a un anillo unífero conmutativo.

Capítulo IV

Polinomios con una o varias variables

Dejando \mathbf{Z} aparte, el anillo $K[X]$ de los polinomios con coeficientes en un cuerpo conmutativo K es el ejemplo más familiar de anillo principal. Las propiedades aritméticas de los §§ 2 y 3 dependen únicamente de esta estructura, de ahí su exacta correspondencia con las de \mathbf{Z} , que se puede extender a todo anillo principal.

§ IV.1 DEFINICIÓN DE $A[X]$, PROPIEDADES GENERALES

DEFINICIÓN IV.1.1

Si A es un anillo conmutativo unífero, se llama **polinomio con coeficientes en A** a una sucesión $(a_0, a_1, \dots, a_n, \dots)$ de elementos de A en la cual sólo un número finito de términos es no nulo. A este polinomio se le llama **normalizado** (o **unitario**) si su último coeficiente no nulo, llamado coeficiente **dominante**, es igual a 1.

Notación. Al conjunto de los polinomios con coeficientes en un anillo A se le designa por $A[X]$. En $A[X]$ se definen dos leyes de composición interna:

Adición. Si $P = (a_0, a_1, \dots, a_n, \dots) = (a_p)_{p \in \mathbf{N}}$ y $Q = (b_0, b_1, \dots, b_n, \dots) = (b_p)_{p \in \mathbf{N}}$,

se define la *suma* $P + Q$ por:

$$P + Q = (a_0 + b_0, a_1 + b_1, \dots, a_n + b_n, \dots) = (a_p + b_p)_{p \in \mathbf{N}}.$$

Multiplicación. Conservamos las anteriores notaciones. Para $m \in \mathbf{N}$ definimos

$$(1) \quad c_m = \sum_{p+q=m} a_p b_q.$$

Puesto que los términos (a_p) y (b_q) son nulos a partir de un cierto lugar, existe un $m_0 \in \mathbf{N}$ tal que $(n \geq m_0) \Rightarrow (a_n = b_n = 0)$. Según (1), para $m \geq 2m_0$, se tiene: $c_m = 0$ (para cada término $a_p b_q$ del segundo miembro de (1), uno de los enteros p, q es $\geq m_0$). Ello muestra que la sucesión $(c_m)_{m \geq 0}$ definida por (1) es un polinomio. Por definición, este polinomio es el *producto* PQ de P y Q .

TEOREMA IV.1.1

|| *Provisto de estas dos leyes de composición, $A[X]$ es un anillo conmutativo unífero.*

Demostración

a) Para la adición, $A[X]$ es un grupo abeliano. En efecto, la asociatividad y la conmutatividad de la adición en $A[X]$ resultan de las correspondientes propiedades de la adición en el grupo aditivo A . El elemento neutro (designado por 0) es el polinomio definido por la sucesión que posee todos los coeficientes nulos, que es el *polinomio nulo*; el opuesto de $P = (a_p)_{p \in \mathbf{N}}$ es $-P = (-a_p)_{p \in \mathbf{N}}$.

b) La conmutatividad, asociatividad y distributividad respecto de la suma, del producto, resultan de estas mismas propiedades en A . Por ejemplo, comprobemos la asociatividad:

Sean $P = (a_p)_{p \in \mathbf{N}}$, $Q = (b_k)_{k \in \mathbf{N}}$, $R = (c_r)_{r \in \mathbf{N}}$ tres polinomios. Se tiene

$$PQ = (\beta_m)_{m \in \mathbf{N}}; \quad (PQ)R = (\gamma_n)_{n \in \mathbf{N}}; \quad \gamma_n = \sum_{k+l=n} \beta_k c_l; \quad \beta_k = \sum_{p+k=k} a_p b_q.$$

En virtud de la asociatividad del producto y de la conmutatividad de la suma en A , se deduce de todo ello:

$$\gamma_n = \sum_{p+q+r=n} a_p b_q c_r.$$

Igualmente se demostraría que $P(QR) = (\gamma_n)_{n \in \mathbf{N}}$. El polinomio $(e_p)_{p \in \mathbf{N}}$ en el que $e_0 = 1$ y $e_p = 0$ para $p \geq 1$ es, evidentemente, elemento neutro para el producto. c.q.d.

DEFINICIÓN IV.1.2

} *Dotado de la estructura de anillo definida anteriormente, el conjunto $A[X]$ se llama **anillo de los polinomios con una variable (o indeterminada)**, con coeficientes en A . Este anillo se designa también por $A[X]$.*

Inclusión de A en $A[X]$. La A -álgebra $A[X]$

Definimos la aplicación $j: A \rightarrow A[X]$ por $j(a) = (a, 0, 0, \dots)$ (sucesión en que el término de índice 0 vale a y los restantes términos son nulos). Es claro que j es un *homomorfismo de anillos inyectivo*. Por esta razón, se identifica A con un subanillo de $A[X]$, omitiendo indicar la aplicación j . A los elementos de A se les llama entonces *constantes de $A[X]$* .

En $A[X]$ se puede definir una estructura de A -módulo y también de A -álgebra: el producto de $a \in A$ por el polinomio $P = (a_p)_{p \in \mathbf{N}}$ es el polinomio $(aa_p)_{p \in \mathbf{N}}$.

Grado

Si $-\infty$ es un símbolo, designaremos por \mathbf{N} al conjunto $\mathbf{N} \cup \{-\infty\}$. Se extiende la relación de orden de \mathbf{N} y la adición de \mathbf{N} a \mathbf{N} por medio de las fórmulas:

$$-\infty \leq n \quad (n \in \mathbf{N}), \quad -\infty + n = -\infty.$$

DEFINICIÓN IV.1.3

- $\}$ El grado del polinomio nulo es $-\infty$.
- $\}$ El grado de un polinomio no nulo $P = (a_0, a_1, \dots, a_n, \dots)$, es el mayor entero k tal que $a_k \neq 0$.
- $\}$ En todos los casos, el grado de un polinomio P se designa por $\text{gr}(P)$.

IV.1.2 Para todos los polinomios $P, Q \in A[X]$ se tiene:

- \parallel (2) $\text{gr}(P + Q) \leq \sup(\text{gr}(P), \text{gr}(Q)).$
- \parallel (3) $\text{gr}(PQ) \leq \text{gr}(P) + \text{gr}(Q).$

Demostración. (2) es evidente. Observemos que, siempre que

$$\text{gr}(P) \neq \text{gr}(Q)$$

se tiene la igualdad en (2). La relación (3) es evidente si $P = 0$ o $Q = 0$. Supongamos que P y Q son no nulos, $P = (a_p)_{p \geq 0}$, $Q = (b_q)_{q \geq 0}$ y pongamos $m = \text{gr}(P)$, y $n = \text{gr}(Q)$. Para $p > m$ (resp. $q > n$) se tiene: $a_p = 0$ (resp. $b_q = 0$). Luego, para $p + q > m + n$, se tiene: $a_p b_q = 0$.

Esto demuestra que el término

$$c_M = \sum_{p+q=M} a_p b_q$$

del polinomio $PQ = (c_n)_{n \geq 0}$ es nulo para $M > m + n$, luego (3) ha quedado establecido. c.q.d.

TEOREMA IV.1.3

Si el anillo A es íntegro (luego en particular, si A es un cuerpo) $A[X]$ es íntegro.
 En otras palabras, la relación $PQ = 0$. ($P, Q \in A[X]$) implica entonces $P = 0$ o $Q = 0$.

Demostración. Utilizamos las mismas notaciones que en la demostración de IV.1.2. El término de índice $m + n$ de PQ es

$$c_{m+n} = a_m b_n.$$

Puesto que $a_m \neq 0$, $b_n \neq 0$, se tiene: $c_{m+n} \neq 0$ (pues A es íntegro). Dicho de otra manera, en este caso la fórmula (3) se convierte en

$$(4) \quad \text{gr}(PQ) = \text{gr}(P) + \text{gr}(Q),$$

y las relaciones $P \neq 0$ y $Q \neq 0$ implican $PQ \neq 0$. c.q.d.

IV.1.3 es un ejemplo de «teorema de permanencia», es decir, de teorema verdadero en $A[X]$ cuando es verdadero en A ; veremos otro ejemplo en el capítulo XIV.

Valoración (u orden)

DEFINICIÓN IV.1.4

La valoración del polinomio nulo es $+\infty$.
 La valoración de un polinomio no nulo $P = (a_0, a_1, \dots, a_n, \dots)$ es el menor entero k tal que $a_k \neq 0$.
 A la valoración de un polinomio P se le llama también el orden de P ; en general se le designa por $\omega(P)$.

Las propiedades de la valoración son análogas a las del grado; nos limitaremos a enunciarlas:

$$\omega(P + Q) \geq \inf(\omega(P), \omega(Q)),$$

$$\omega(PQ) \geq \omega(P) + \omega(Q);$$

y, si el anillo de base es íntegro:

$$\omega(PQ) = \omega(P) + \omega(Q).$$

Generador de $A[X]$

Sea X el polinomio $(e_p)_{p \in \mathbb{N}}$ definido por $e_1 = 1$, y $e_p = 0$ para $p \neq 1$, o $X = (0, 1, 0, \dots, 0, \dots)$: a X se le llama la variable (o indeterminada). Para cada $n \in \mathbb{N}$, el polinomio $X^n = (f_p)_{p \in \mathbb{N}}$ está definido por: $f_n = 1$, $f_p = 0$ para $p \neq n$.

Consideremos un polinomio cualquiera $P = (a_n)_{n \in \mathbb{N}}$; puesto que los a_n son nulos a partir de un cierto lugar, la suma $\sum_{n \in \mathbb{N}} a_n X^n$ tiene sentido. Y, en virtud de lo precedente, esta suma es precisamente igual a P . En consecuencia, se puede enunciar:

IV.1.4 Si X es la variable, todo polinomio $P \in A[X]$ se escribe de una manera
 \parallel y sólo de una en la forma $\sum_{n \in \mathbb{N}} a_n X^n$, en donde los a_n son los coeficientes de P .

Se expresa IV.1.4 diciendo que X es un generador de la A -álgebra $A[X]$. Según el capítulo III, § 9, IV.1.4 expone también el hecho de que el conjunto de los polinomios $(X^n)_{n \in \mathbb{N}}$ forma una base del A -módulo $A[X]$.

Si n designa el grado de $P = \sum_{p \in \mathbb{N}} a_p X^p$, se puede escribir igualmente:

$$P = \sum_{k=0}^n a_k X^k \quad (\text{ordenación de } P \text{ según las potencias crecientes}),$$

$$P = \sum_{k=n}^0 a_k X^k \quad (\text{ordenación de } P \text{ según las potencias decrecientes}).$$

Substitución de un polinomio en otro

DEFINICIÓN IV.1.5

Sean $P = \sum_{k=0}^n a_k X^k$ y $Q = \sum_{k=0}^p b_k X^k$ dos polinomios.
 Se llama **compuesto** de P y Q (notación $P \circ Q$ o $P(Q)$) al polinomio $\sum_{k=0}^n a_k Q^k$. Se dice también que $P \circ Q$ se deduce de P substituyendo X por Q .

Propiedades de la operación $(P, Q) \mapsto P \circ Q$:

$$P \circ X = P,$$

$$(P_1 + P_2) \circ Q = P_1 \circ Q + P_2 \circ Q;$$

pero, en general,

$$P \circ (Q_1 + Q_2) \neq P \circ Q_1 + P \circ Q_2.$$

Por ejemplo, si $P = 1 + X^3$, $Q_1 = X^2$, $Q_2 = X$, se tiene:

$$P \circ (Q_1 + Q_2) = 1 + (X + X^2)^3, \quad \text{y} \quad P \circ Q_1 + P \circ Q_2 = 2 + X^3 + X^6.$$

Elementos invertibles de $A[X]$

TEOREMA IV.1.5

|| Si el anillo A es íntegro, el grupo de las unidades de $A[X]$ coincide con el grupo de las unidades de A .

Demostración. Si $P \in A[X]$, $Q \in A[X]$ y $PQ = 1$, se tiene ante todo $P \neq 0$ y $Q \neq 0$, luego, en virtud de (4):

$$\text{gr}(P) + \text{gr}(Q) = \text{gr}(1) = 0;$$

esta última relación implica que $\text{gr}(P) = \text{gr}(Q) = 0$, luego P y Q son constantes invertibles. El recíproco es evidente. c.q.d.

Si A no es íntegro, IV.1.5 no es verdadero: por ejemplo, si existe $a \in A$ tal que $a \neq 0$ y $a^2 = 0$, se tiene:

$$(aX + 1)(-aX + 1) = 1.$$

Observemos, finalmente, que si A es íntegro, según IV.1.5, los polinomios de grado > 0 no son invertibles.

Cambio del anillo de base

$\rho : A \rightarrow B$ designa a un homomorfismo no nulo de anillos uníferos conmutativos. Hagamos corresponder, a todo $P \in A[X]$ definido por $P = \sum a_n X^n$, el polinomio $Q \in B[X]$ definido por: $Q = \sum \rho(a_n) X^n$.

La aplicación $P \mapsto Q$ es un homomorfismo de anillos

$$\bar{\rho} : A[X] \rightarrow B[X],$$

llamado *extensión* de ρ . Si ρ es inyectivo (resp. epiyectivo), $\bar{\rho}$ es inyectivo (resp. epiyectivo).

Este razonamiento se aplica especialmente cuando A es un subanillo de B y $\rho : A \rightarrow B$ es la inyección canónica. En este caso, $\bar{\rho}$ permite identificar $A[X]$ con un subanillo de $B[X]$. Pero entonces es preciso distinguir perfectamente bien entre las propiedades de un polinomio $P \in A[X]$, según que lo consideremos como elemento de $A[X]$ o de $B[X]$; (cf., por ejemplo, la noción de irreducibilidad: ver § 3).

§ IV.2 DIVISIÓN EUCLÍDEA.

PROPIEDADES ARITMÉTICAS DE $K[X]$

CUANDO K ES UN CUERPO CONMUTATIVO

División euclídea

- De momento, K designa a un anillo conmutativo unífero (en adelante, supondremos que K es un cuerpo).

TEOREMA IV.2.1

Sea A un polinomio cualquiera de $K[X]$ y B un polinomio, tcl que el coeficiente de su término de más alto grado sea invertible en K . Existen polinomios Q y R , unívocamente determinados, tales que

$$\begin{array}{ll} (1) & A = BQ + R, \\ (2) & \text{gr}(R) < \text{gr}(B); \end{array}$$

a Q se le llama *cociente*, y a R , *resto* en la división euclídea de A por B .

Demostración

a) *Unicidad*. Si las parejas de polinomios (Q, R) , (Q', R') satisfacen (1) y (2), se tiene, por sustracción miembro a miembro,

$$(3) \quad B(Q' - Q) = R - R'.$$

El coeficiente del término de más alto grado de B es invertible, por lo tanto regular en K . Volviendo a la demostración de IV.1.3 la relación (3) implica,

$$\text{gr}[B(Q' - Q)] = \text{gr}(B) + \text{gr}(Q' - Q);$$

si $R \neq R'$, se deduce que $Q' - Q \neq 0$, de donde $\text{gr}[B(Q' - Q)] \geq \text{gr}(B)$; pero por otro lado,

$$\text{gr}[B(Q' - Q)] = \text{gr}(R - R') \leq \sup[\text{gr}(R), \text{gr}(R')] < \text{gr}(B),$$

de donde se sigue una contradicción. Luego $R = R'$. Finalmente se establece que $Q = Q'$ teniendo en cuenta que B no es un divisor de cero en $K[X]$ (el coeficiente de su término de más alto grado es invertible).

b) *Existencia.* Si $A = 0$, basta hacer $Q = R = 0$. Si no, se razona por recurrencia sobre $n = \text{gr}(A)$. Escribimos $B = b_p X^p + \dots + b_0$, con $p = \text{gr}(B)$. Por hipótesis, b_p es invertible en K . Cuando $n < p$, se hace $Q = 0$, $R = A$. Supongamos la propiedad verdadera para todos los polinomios A de grado $\leq n$, y demostrémosla para

$$A = a_{n+1} X^{n+1} + \dots + a_0, \quad (n+1 = \text{gr}(A), \quad a_{n+1} \neq 0).$$

El polinomio $A_1 = A - BQ_1$, en donde $Q_1 = \frac{a_{n+1}}{b_p} X^{n+1-p}$, es de grado $\leq n$.

Por la hipótesis de recurrencia, existen dos polinomios Q_2 y R_2 , con $\text{gr}(R_2) < p$, tales que $A_1 = BQ_2 + R_2$. Se tiene:

$$A = A_1 + BQ_1 = B(Q_1 + Q_2) + R_2.$$

Por lo tanto es suficiente hacer $R = R_2$, $Q = Q_1 + Q_2$ a fin de obtener (1) y (2). c.q.d.

La disposición práctica clásica de la división se inspira en el anterior razonamiento por recurrencia.

Ejemplo

$$\begin{array}{rcl} A = X^5 + 2X^3 - 3X - 2, & B = X^3 + X + 1 & \\ X^5 & + 2X^3 & - 3X - 2 \quad \Big| \quad X^3 + X + 1 \\ A_1 = & X^3 - X^2 - 3X - 2 & \Big| \quad X^2 + 1 \\ A_2 = & -X^2 - 4X - 3 & \Big| \quad Q_1 \quad Q_2 \end{array} \quad \begin{array}{l} Q = X^2 + 1 \\ R = -X^2 - 4X - 3. \end{array}$$

Cuando K es un *cuerpo conmutativo*, cualquier pareja de polinomios (A, B) , $B \neq 0$ verifica las hipótesis del teorema IV.2.1.

- Por este motivo, en lo que resta de este capítulo, **supondremos que el anillo de base K es un cuerpo conmutativo.**

Recordemos algunas definiciones acerca de los divisores (dadas ya en el capítulo III en un cuadro más general). Se dice que $B \in K[X]$ es un *divisor* de A , o que A es un *múltiplo* de B , si existe $Q \in K[X]$ tal que $A = BQ$. Para ello, es necesario y suficiente que, en la división euclídea de A por B , el resto sea nulo. (Cuando $B \neq 0$.)

Dos polinomios A y B están *asociados* si cada uno de ellos es múltiplo del otro. Ello equivale a decir que existe $\lambda \in K^*$ tal que $B = \lambda A$, o que $A = B = 0$. Diremos también que tales polinomios *son proporcionales*, o que *difieren únicamente en un factor constante*.

Ideales de $K[X]$

En lo que sigue designaremos por (A_1, A_2, \dots, A_n) al ideal de $K[X]$ engendrado por los polinomios A_1, \dots, A_n . Recordemos que (A_1, \dots, A_n) es el conjunto de los polinomios A de la forma $A = \sum_{i=1}^n U_i A_i$, en donde los U_i son polinomios arbitrarios de $K[X]$.

En particular, el ideal principal engendrado por A se designará por (A) .

TEOREMA IV.2.2

Si K es un cuerpo conmutativo, todo ideal de $K[X]$ es principal.
(En otras palabras (cf. Cap. III) $K[X]$ es un **anillo principal**). Preciando, a todo ideal \mathfrak{g} de $K[X]$, se le puede asociar un polinomio P , único salvo un factor no nulo, tal que $\mathfrak{g} = (P)$.

Demostración. Si $\mathfrak{g} = \{0\}$ se puede tomar $P = 0$ y ésta es la única solución. Si $\mathfrak{g} \neq \{0\}$, sea $\mathcal{G} = \{\text{gr}(A) \mid A \in \mathfrak{g} \text{ y } A \neq 0\}$. El conjunto \mathcal{G} es una parte no vacía de \mathbf{N} ; existe, pues, un primer elemento $d \geq 0$. Sea $P \in \mathcal{G}$ tal que $\text{gr}(P) = d$. Se verifica, evidentemente, $(P) \subset \mathfrak{g}$. Por otro lado, si $A \in \mathfrak{g}$, la división euclídea de A por P da $A = PQ + R$, con $\text{gr}(R) < d$, y ($A \in \mathfrak{g}$ y $PQ \in \mathfrak{g}$) implica $A - PQ = R \in \mathfrak{g}$. Es, pues, necesario que $\text{gr}(R) = -\infty$, luego $R = 0$ lo que demuestra que A es un múltiplo de P ; de ahí la inclusión $\mathfrak{g} \subset (P)$, o sea, en fin, $\mathfrak{g} = (P)$ ya que $P \in \mathfrak{g}$.

Si $\mathfrak{g} = (P_1) = (P_2)$, con $\mathfrak{g} \neq \{0\}$, P_1 y P_2 son no nulos y múltiplos el uno del otro, luego están asociados. c.q.d.

Máximo común divisor

Sea A_1, A_2, \dots, A_n una familia finita de polinomios *no todos nulos* y sea

$$J = (A_1, A_2, \dots, A_n)$$

el ideal engendrado por estos polinomios. Según IV.2.2 existe un polinomio D , único salvo un factor no nulo, tal que

$$J = (A_1, A_2, \dots, A_n) = (D).$$

Este polinomio D es un divisor común a todos los A_i (ya que $A_i \in J$); y todo polinomio que divida a cada uno de los A_i divide a D (ya que $D \in J$):

DEFINICIÓN IV.2.1

Si los polinomios A_1, A_2, \dots, A_n son todos no nulos, el polinomio D (único salvo un factor) tal que $(A_1, A_2, \dots, A_n) = (D)$ se llama **máximo común divisor** (mcd) de los polinomios

$$A_1, A_2, \dots, A_n.$$

En lo que sigue, al mcd de A_1, A_2, \dots, A_n lo designaremos por

$$A_1 \top A_2 \top \dots \top A_n$$

(no existe una notación universal).

$D = A_1 \top A_2 \top \dots \top A_n$ se caracteriza por las relaciones:

- D divide a cada uno de los polinomios A_i ($i = 1, 2, \dots, n$),
- todo divisor común a los A_i divide a D .

DEFINICIÓN IV.2.2

Se dice que los polinomios A_1, A_2, \dots, A_n son **primos entre sí** en conjunto si su mcd es una constante no nula (que podemos suponer igual a 1), en otras palabras, si carecen de divisor común de grado > 0 .

TEOREMA IV.2.3 (Bezout)

Para que los polinomios A_1, A_2, \dots, A_n sean primos entre sí es necesario y suficiente que existan polinomios U_1, U_2, \dots, U_n tales que

$$(4) \quad \sum_{i=1}^n U_i A_i = 1.$$

Demostración. La condición es necesaria, puesto que si (A_1, A_2, \dots, A_n) son primos entre sí, $1 \in (A_1, A_2, \dots, A_n)$ lo que implica la existencia de U_1, \dots, U_n tales que se verifique (4).

Recíprocamente, si (4) se verifica, todo polinomio que divida a A_1, A_2, \dots, A_n divide también a $\sum_{i=1}^n U_i A_i$, luego divide a 1. c.q.d.

Propiedades del mcd

Por la misma definición, la operación $(A, B) \mapsto A \top B$ es conmutativa; de ello se deduce que, para toda permutación $\sigma \in \mathfrak{S}_n$, se verifica:

$$A_1 \top A_2 \top \dots \top A_n = A_{\sigma(1)} \top A_{\sigma(2)} \top \dots \top A_{\sigma(n)}.$$

IV.2.4 (*Asociatividad*). Para toda terna de polinomios A, B, C , se tiene:

$$\parallel A \top B \top C = (A \top B) \top C = A \top (B \top C).$$

Resulta de las definiciones.

Se deduce, en particular, la regla siguiente (cf. Cap. II, § 2).

Para determinar el mcd de n polinomios, se pueden substituir p cualesquiera de ellos por su mcd ($p < n$).

IV.2.5 Para todo polinomio A_1, \dots, A_n, B , se tiene:

$$\parallel (5) \quad B.(A_1 \top A_2 \top \dots \top A_n) = (BA_1) \top (BA_2) \top \dots \top (BA_n);$$

(distributividad del producto respecto de \top).

La demostración es inmediata.

TEOREMA IV.2.6 (T. de Gauss)

$$\parallel \text{Si } A, B, C \text{ son tres polinomios, si } C \text{ es primo con } B \text{ y divide a } AB, \text{ divide a } A.$$

Primera demostración. Según IV.2.5, $B \top C = 1$ implica

$$(AB) \top (AC) = A.$$

Y puesto que C divide tanto a AB como a AC , divide a A .

Segunda demostración. Sea L tal que $AB = CL$, y U, V tales que $UB + VC = 1$. Se deduce que $UAB + VAC = A$, o sea

$$ULC + VC = A, \quad (UL + AV)C = A,$$

lo que demuestra que C divide a A .]]

Generalización (inmediata). Si el polinomio C divide al producto $A_1 \cdot A_2 \cdot \dots \cdot A_n$ y es primo con cada uno de los polinomios A_1, A_2, \dots, A_{n-1} , divide a A_n (demostración por recurrencia a partir de IV.2.6).

TEOREMA IV.2.7

$$\parallel \text{Sean } A, B, C \in K[X]; \text{ si } A \text{ y } B \text{ son primos entre sí y dividen a } C, \text{ entonces } C \text{ es múltiplo de } AB.$$

Demostración. Existen polinomios P, Q, U, V tales que $C = AP = BQ$, $UA + VB = 1$. Se deduce que: $UAC + VBC = C$, o sea

$$UQAB + VPAB = C, \quad (UQ + VP)AB = C,$$

luego AB divide a C . c.q.d.

TEOREMA IV.2.8

|| Si A es primo con B , y primo con C , es primo con el producto BC .

Demostración (abreviada). $UA + VB = 1$ y $U'A + V'C = 1$ implican por multiplicación $(UU'A + UV'C + VU'B)A + (VV')BC = 1$. c.q.d.

Mínimo común múltiplo

Sea A_1, \dots, A_n una familia finita de polinomios; la intersección

$$(A_1) \cap (A_2) \cap \dots \cap (A_n)$$

de ideales (A_i) es un ideal de $K[X]$, formado por los múltiplos comunes a A_1, A_2, \dots, A_n .

Si M es el polinomio (único salvo un factor) de $K[X]$ tal que $\bigcap_{i=1}^n (A_i) = (M)$, se dice que M es el *mínimo común múltiplo* (mcm) de los polinomios (A_i) . Lo designaremos por:

$$M = A_1 \perp A_2 \perp \dots \perp A_n.$$

Está caracterizado por las dos condiciones:

- es un múltiplo de A_1, A_2, \dots, A_n ;
- todo múltiplo común a A_1, A_2, \dots, A_n es un múltiplo de M .

Las propiedades inmediatas del mcm son las propiedades *conmutativa, asociativa y distributiva del producto*, que implican, respectivamente, las fórmulas siguientes:

$$A_1 \perp A_2 \perp \dots \perp A_n = A_{\sigma(1)} \perp A_{\sigma(2)} \perp \dots \perp A_{\sigma(n)} \quad (\sigma \in \mathfrak{S}_n)$$

$$A \perp B \perp C = (A \perp B) \perp C = A \perp (B \perp C),$$

$$B.(A_1 \perp A_2 \perp \dots \perp A_n) = (BA_1) \perp (BA_2) \perp \dots \perp (BA_n).$$

En el caso de dos polinomios, se tiene además la propiedad siguiente:

TEOREMA IV.2.9

$$\left\| \begin{array}{l} \text{Sean } A, B \text{ dos polinomios no nulos, } D = A \top B \text{ y } M = A \perp B. \text{ Se tiene:} \\ M = \frac{AB}{D}. \end{array} \right.$$

(Esta propiedad no se generaliza al caso de n polinomios.)

Demostración. Hagamos $A = DA_1$ y $B = DB_1$; $AB/D = A_1 B_1 D$ es un múltiplo común a A y B . Recíprocamente, si N es un múltiplo de A y de B , es un múltiplo de D , y $P = \frac{N}{D}$ es un múltiplo común a $A_1 B_1$. Los polinomios A_1, B_1 son primos entre sí, P es un múltiplo del producto $A_1 B_1$ (T. IV.2.7), y N es múltiplo de $A_1 B_1 D = \frac{AB}{D}$. c.q.d.

Interpretación con la ayuda del conjunto ordenado de los ideales

Cuando se ordena por inclusión el conjunto de los ideales de $K[X]$, el ideal $(A_1 \perp A_2 \perp \dots \perp A_n)$ es el *ínfimo* de los ideales (A_i) , y el ideal $(A_1 \top A_2 \top \dots \top A_n)$ es el *supremo* de los ideales (A_i) (cf. Cap. III, § 5).

*** Extensión de escalares**

Sean K, L cuerpos, sea ρ un isomorfismo de K en un subcuerpo de L . Sean $A, B \in K[X]$ y $D = \text{mcd}(A, B)$. Llamemos $\bar{\rho} : K[X] \rightarrow L[X]$ a la extensión de ρ (ver § 1). Entonces $\bar{\rho}(D) = \text{mcd}(\bar{\rho}(A), \bar{\rho}(B))$. En efecto, $\bar{\rho}(D)$ divide a $\bar{\rho}(A)$ y $\bar{\rho}(B)$, y existen U y $V \in K[X]$ tales que $UA + VB = D$, de donde

$$\bar{\rho}(U) \bar{\rho}(A) + \bar{\rho}(V) \bar{\rho}(B) = \bar{\rho}(D),$$

lo que prueba que todo divisor común a $\bar{\rho}(A)$ y $\bar{\rho}(B)$ debe dividir a $\bar{\rho}(D)$.

Todo ello se aplica, en particular, al caso en que ρ es la inyección canónica $K \rightarrow L$, cuando K es un subcuerpo de L . No debe preocuparnos saber sobre qué cuerpo de base (K o L) nos hallamos para hablar del mcd y del mcm, de divisiones euclideas, de polinomios primos entre sí, etc.

Más adelante veremos que no sucede lo mismo para otras propiedades aritméticas.

§ IV.3 ALGORITMO DE EUCLIDES

La teoría precedente no proporciona ningún método práctico para determinar el mcd de una familia de polinomios, ni tampoco para calcular los polinomios U_i

cuya existencia se halla establecida por el teorema de Bezout. Por todo ello el algoritmo de Euclides, que se describe a continuación, posee gran interés práctico.

Se observará que este algoritmo sólo trata del caso de dos polinomios. Pero el teorema de asociatividad IV.2.4 permite transformar el caso general al caso de dos polinomios.

Además, este algoritmo es independiente de la teoría del § 2, y nos proporciona una nueva demostración de la existencia del mcd.

Construcción

Sean A, B dos polinomios, $B \neq 0$ y busquemos $D = A \top B$. Definimos por recurrencia polinomios $(R_k)_{k \geq 0}$ y $(Q_k)_{k \geq 1}$ por medio de las condiciones siguientes:

$$\begin{aligned} R_0 &= A, \quad R_1 = B, \quad \text{y, para } k \geq 1 : \\ a) \text{ si } R_k &\neq 0 \quad (1) \quad \left. \begin{aligned} R_{k-1} &= R_k Q_k + R_{k+1} \\ \text{gr}(R_{k+1}) &< \text{gr } R_k \end{aligned} \right\} \begin{array}{l} \text{división euclídea de } R_{k-1} \text{ por} \\ R_k; \end{array} \\ b) \text{ si } R_k &= 0 \quad R_{k+p} = 0 \text{ para } p \geq 1, \quad Q_{k+p} = 0 \text{ para } p \geq 0. \end{aligned}$$

Si se tiene $R_k \neq 0$ para todo k , la sucesión infinita de enteros positivos $\text{gr}(R_k)$, sería estrictamente decreciente, lo cual es absurdo. Luego existe $k \geq 2$ tal que $R_k = 0$. Sea k_0 el menor de estos enteros, luego por construcción, $R_1, R_2, \dots, R_{k_0-1}$ son $\neq 0$.

DEFINICIÓN IV.3.1

$\left\{ \begin{array}{l} \text{Con las notaciones anteriores, la sucesión finita de relaciones (1) para} \\ k \leq k_0 - 1 \text{ es la sucesión de } \mathbf{divisiones sucesivas} \text{ de } A \text{ por } B, \text{ la} \\ \text{sucesión } R_2, \dots, R_{k_0-1} \text{ es la sucesión de } \mathbf{restos sucesivos}. \end{array} \right.$

R_{k_0-1} es, pues, el último resto no nulo.

IV.3.1 El mcd de A y B es el último resto no nulo en la sucesión de divisiones
|| sucesivas de A por B .

Demostración. Utilizamos las notaciones que preceden a la definición IV.3.1. Para $k \leq k_0 - 1$,

$$R_{k-1} = R_k Q_k + R_{k+1}, \quad \text{gr}(R_{k+1}) < \text{gr}(R_k).$$

El conjunto de divisores comunes a R_{k-1} y R_k coincide con el conjunto de divisores comunes a R_k y R_{k+1} : en el lenguaje de los ideales:

$$(R_{k-1}, R_k) = (R_k, R_{k+1}).$$

Por recurrencia finita se deducen las igualdades entre ideales:

$$(A, B) = (R_0, R_1) = \cdots = (R_{k_0-1}, R_{k_0}) = (R_{k_0-1}, 0) = (R_{k_0-1}). \text{ c.q.d.}$$

Ejemplos

1) Hallar $D = \text{mcd}(A, B)$, con

$$A = X^5 + X^4 + 2X^3 - 2X + 3, \quad B = X^4 + 3X^3 + 7X^2 + 8X + 6.$$

$$A = BQ_1 + R_2, \quad Q_1 = X - 2, \quad R_2 = X^3 + 6X^2 + 8X + 15,$$

$$B = R_2Q_2 + R_3, \quad Q_2 = X - 3, \quad R_3 = 17X^2 + 17X + 51,$$

$$R_2 = R_3Q_3 + R_4, \quad Q_3 = X + 5, \quad R_4 = 0.$$

Luego $D = R_3 = 17(X^2 + X + 3)$; $\text{mcd}(A, B) = X^2 + X + 3$.

2) Hallar $D = \text{mcd}(A, B)$, $A = X^m - 1$, $B = X^n - 1$. Efectuemos la división euclídea de m por n : $m = nq + r$, $r < n$.

$$A = X^m - 1 = X^{nq+r} - 1 = X^r(X^{nq} - 1) + X^r - 1, \quad \text{d}^0(X^r - 1) < n.$$

$$X^{nq} - 1 = (X^n - 1)(X^{n(q-1)} + X^{n(q-2)} + \cdots + 1) = B \cdot H$$

(en donde H designa a un polinomio). En consecuencia $A = HX^r \cdot B + X^r - 1$.

Luego, si se escriben las divisiones sucesivas de A por B , la sucesión R_k de los restos sucesivos viene dada por $R_k = X^{r_k} - 1$, en donde (r_k) es la sucesión de los restos en las divisiones sucesivas de m por n en \mathbf{N} . Se deduce también que $D = X^d - 1$, en donde $d = \text{mcd}(m, n)$. c.q.d.

El algoritmo de Euclides nos permitirá precisar el teorema de Bezout (e incluso suministrar una nueva demostración) en el caso de dos polinomios.

IV.3.2 Si los polinomios A, B son primos entre sí y no constantes los dos, existe

un par único de polinomios (U, V) tal que

$$(2) \quad UA + VB = 1, \quad \text{gr}(U) < \text{gr}(B), \quad \text{gr}(V) < \text{gr}(A).$$

Demostración

a) *Unicidad.* Si (U_1, V_1) y (U_2, V_2) son dos pares de polinomios que verifican (2), se tiene $(U_1 - U_2)A = (V_2 - V_1)B$. Según el teorema de Gauss, si $U_1 - U_2$

fuera no nulo, A dividiría a $V_2 - V_1$, lo cual no es posible puesto que $\text{gr}(V_2 - V_1) < \text{gr}(A)$. Luego $U_1 = U_2$ y $V_1 = V_2$.

b) *Existencia.* Si se conoce un par (U_0, V_0) de polinomios que satisfaga $U_0 A + V_0 B = 1$ y si $\text{gr}(U_0) \geq \text{gr}(B)$, es suficiente hacer la división euclídea de U_0 por B , $U_0 = BQ + U$, con $\text{gr}(U) < \text{gr}(B)$ y escribir $V = V_0 + AQ$ para obtener un par (U, V) de polinomios que satisfagan $UA + VB = 1$ y $\text{gr}(U) < \text{gr}(B)$, y, por consiguiente (ya que A y B no son ambos constantes), $\text{gr} V < \text{gr} A$, es decir (2). La existencia de (U, V) resulta, pues, del teorema IV.2.3. c.q.d.

Las divisiones sucesivas de A por B nos permiten obtener directamente el par (U, V) sin necesidad de utilizar el teorema IV.2.3.

Consideremos de nuevo las relaciones (1). R_{k_0-1} es el último resto no nulo, luego (puesto que A y B son primos entre sí) podemos suponer que $R_{k_0-1} = 1$. Probaremos, por recurrencia, sobre el entero $p = k_0 - k$, que existe una sucesión de polinomios S_k tal que se tiene

$$(3) \quad S_{k+1} R_k + S_k R_{k+1} = 1 \quad \text{y} \quad \text{gr}(S_k) < \text{gr}(R_k)$$

para $k = k_0, \dots, 1$.

En efecto, la relación (3) es verdadera cuando $p = k_0 - k = 2$ con

$$S_{k_0-2} = 1, \quad S_{k_0-1} = 0.$$

Supongámosla verdadera para $k_0 - k \leq p$, y probemos que es también verdadera para $k_0 - k = p + 1$. Reemplazando en (3) R_{k+1} por $R_{k-1} - R_k Q_k$ (relación (1)) se tiene en efecto:

$$S_{k+1} R_k + S_k(R_{k-1} - R_k Q_k) = 1,$$

o sea

$$S_k R_{k-1} + S_{k-1} R_k = 1,$$

poniendo

$$(4) \quad S_{k-1} = S_{k+1} - S_k Q_k;$$

por la hipótesis de recurrencia, se tienen las desigualdades

$$\text{gr}(S_{k+1}) < \text{gr}(R_{k+1}), \quad \text{gr}(S_k) < \text{gr}(R_k),$$

de las que se deduce

$$\text{gr}(S_{k-1}) < \sup[\text{gr}(R_{k-1}), \text{gr}(R_k Q_k)].$$

Pero la identidad de división $R_{k-1} = R_k Q_k + R_{k+1}$ nos muestra que se verifica

$$\text{gr}(R_k Q_k) = \text{gr}(R_{k-1}) \quad \text{y} \quad \text{gr}(R_{k+1}) < \text{gr}(R_{k-1})$$

de donde se deduce

$$\text{gr}(S_{k-1}) < \text{gr}(R_{k-1}),$$

lo que demuestra que (3) es verdadera si se reemplaza k por $k-1$ (lo que equivale a reemplazar p por $p+1$). Paso a paso vemos que (3) es verdadera para $k = k_0 - 2, k_0 - 3, \dots, 1, 0$; y para $k = 0$ se obtienen (teniendo en cuenta que $R_0 = A$ y $R_1 = B$) dos polinomios S_0, S_1 que satisfacen

$$S_1 A + S_0 B = 1, \quad \text{gr}(S_0) < \text{gr}(A), \quad \text{gr}(S_1) < \text{gr}(B).$$

Los polinomios buscados son, pues, $U = S_1$ y $V = S_0$.

En la práctica se partirá de la última división escrita con resto no nulo, o sea $R_{k_0-3} = R_{k_0-2} Q_{k_0-2} + 1$ (si $R_{k_0-1} = 1$); se substituirá R_{k_0-2} por su expresión extraída de la división precedente, y así sucesivamente. En cada identidad obtenida se reemplaza el resto R_k de índice mayor que en ella interviene por $R_{k-2} - R_{k-1} Q_{k-1}$; finalmente se obtienen los polinomios U, V buscados.

Ejemplos

1) $A = X^7 - X - 1, B = X^5 + 1$, son primos entre sí.

En este caso se tiene:

$$\begin{aligned} A &= BQ_1 + R_2, & Q_1 &= X^2, & R_2 &= -X^2 - X - 1; \\ B &= R_2 Q_2 + R_3, & Q_2 &= -X^3 + X^2 - 1, & R_3 &= -X; \\ R_2 &= R_3 Q_3 + R_4, & Q_3 &= X + 1, & R_4 &= -1; \end{aligned}$$

de donde:

$$1 = R_3 Q_3 - R_2 = (B - R_2 Q_2) Q_3 - R_2, \quad \text{con } R_2 = A - BQ_1,$$

o sea

$$1 = Q_3(B - Q_2(A - BQ_1)) - A + BQ_1 = A(-Q_2 Q_3 - 1) + B(Q_1 + Q_3 + Q_1 Q_2 Q_3).$$

Se obtienen:

$$U = -1 - Q_2 Q_3 = X^4 - X^2 + X, \quad V = Q_1 + Q_3 + Q_1 Q_2 Q_3 = -X^6 + X^4 - X^3 + X + 1.$$

2) $A = X^m$ y $B = (1 - X)^n$ son primos entre sí.

En:

$$1 = [X + (1 - X)]^{m+n-1} = \sum_0^{m+n-1} \binom{m+n-1}{k} X^k (1 - X)^{m+n-1-k}$$

agrupamos los términos de exponente $k \geq m$; se obtiene

$$1 = UA + VB, \text{ con } U = \sum_{0 \leq p \leq m+n-1} \binom{m+n-1}{m+p} X^p (1-X)^{n-1-p},$$

$$V = \sum_{k < m} \binom{m+n-1}{k} X^k (1-X)^{m-1-k}$$

y se tiene además $\text{gr}(U) < n$, $\text{gr}(V) < m$.

§ IV.4 POLINOMIOS IRREDUCIBLES (SOBRE UN CUERPO)

DEFINICIÓN IV.4.1

Un polinomio P es **irreducible** si:

- a) $\text{gr}(P) \geq 1$,
- b) los únicos divisores de P son (salvo factores no nulos) 1 y P .

IV.4.1 Sea P un polinomio irreducible y A un polinomio cualquiera. Si P no divide a A , A y P son primos entre sí.

Demostración. Pongamos $D = P \top A$, puesto que D divide a P , se tiene:

$$D = 1 \text{ o } D = P, \text{ c.q.d.}$$

IV.4.2 Si el polinomio irreducible P divide al producto $A_1 \dots A_n$, divide a uno de los factores A_i .

Demostración. Si P no divide a ninguno de los factores A_i , es primo con cada A_i según IV.4.1, luego es primo con $A_1 A_2 \dots A_n$ (T. IV.2.8). c.q.d.

TEOREMA IV.4.3

Si K es un cuerpo conmutativo, todo polinomio $A \in K[X]$, de grado ≥ 1 , posee un divisor irreducible.

Demostración. El conjunto \mathcal{D} de los polinomios de grado ≥ 1 , que dividen a A , es no vacío puesto que contiene a A ; existe, pues, un $P \in \mathcal{D}$ de grado mínimo. Si P no fuese irreducible, se tendría

$$(1) \quad P = QR, \quad \text{gr}(Q) \geq 1, \quad \text{gr}(R) \geq 1.$$

Puesto que (1) implica $\text{gr}(P) = \text{gr}(Q) + \text{gr}(R)$, se tendría

$$1 \leq \text{gr}(Q) < \text{gr}(P),$$

y Q dividiría a A (puesto que Q divide a P). Luego P no sería de grado mínimo en \mathcal{D} . Resulta, pues, que P es irreducible. c.q.d.

A fin de enunciar correctamente el teorema de factorización, introducimos la noción de *conjunto representativo de polinomios irreducibles*: por definición, un tal conjunto \mathcal{D} es un conjunto de polinomios de $K[X]$ que verifica:

a) todo $P \in \mathcal{D}$ es irreducible;

b) si $P \in \mathcal{D}$, $Q \in \mathcal{D}$ y $P \neq Q$, entonces P y Q no son proporcionales (luego son primos entre sí);

c) todo polinomio irreducible $R \in K[X]$ es tal que existe un $P \in \mathcal{D}$ con P y R proporcionales.

Por ejemplo, el conjunto \mathcal{D}_N de los polinomios *irreducibles normalizados* es representativo.

Se tiene entonces:

TEOREMA IV.4.4

Sea \mathcal{D} un conjunto representativo de polinomios irreducibles de $K[X]$ (en donde K es un cuerpo conmutativo).

Para todo polinomio A **no nulo**, $A \in K[X]$, existe una familia única $(\alpha_P)_{P \in \mathcal{D}}$ de enteros casi todos nulos y un elemento único $u \in K^*$ tales que

$$A = u \prod_{P \in \mathcal{D}} P^{\alpha_P}.$$

Demostración

a) *Unicidad*: Sean $(\alpha_P)_{P \in \mathcal{D}}$, $(\beta_P)_{P \in \mathcal{D}}$, dos familias de enteros casi todos nulos, y $u_1, u_2 \in K^*$ tales que

$$A = u_1 \prod_{P \in \mathcal{D}} P^{\alpha_P} = u_2 \prod_{P \in \mathcal{D}} P^{\beta_P}.$$

Supongamos que exista un $P_0 \in \mathcal{D}$ con $\alpha_{P_0} > \beta_{P_0}$. Se tendría entonces

$$u_1 P_0^{\alpha_{P_0} - \beta_{P_0}} \prod_{P \in \mathcal{D}, P \neq P_0} P^{\alpha_P} = u_2 \prod_{P \in \mathcal{D}, P \neq P_0} P^{\beta_P} = B.$$

Luego el polinomio P_0 debería dividir a uno de los $(P^{\beta_P})_{P \in \mathcal{D}, P \neq P_0}$, puesto que es irreducible. Luego debería dividir a uno de los P ($P \in \mathcal{D}$, $P \neq P_0$), ya que la

hipótesis $\alpha_{P_0} - \beta_{P_0} > 0$ implica que el grado de B es ≥ 1 . Ello es contradictorio, por lo que se deduce que $\alpha_P \leq \beta_P$ para todo $P \in \mathcal{P}$. Igualmente se verá que $\alpha_P \geq \beta_P$ para todo $P \in \mathcal{P}$. Luego $\alpha_P = \beta_P$ para todo $P \in \mathcal{P}$, y puesto que $A \neq 0$, esto implica $u_1 = u_2$.

b) *Existencia*: Para todo $P \in \mathcal{P}$, sea α_P el mayor de los enteros k tales que P^k divide a A ⁽¹⁾. Para toda parte finita J de \mathcal{P} , el polinomio $\prod_{P \in J} P^{\alpha_P}$ divide a A (pues los P^{α_P} son primos entre sí), de donde $\sum_{P \in J} \alpha_P \leq \text{gr}(A)$: luego los α_P son casi todos nulos, y este razonamiento demuestra además que el polinomio $C = \prod_{P \in \mathcal{P}} P^{\alpha_P}$ divide a A . Por definición de los α_P , el polinomio A/C no puede poseer ningún divisor irreducible. Luego (T. IV.4.3) $\frac{A}{C} \in K^*$. c.q.d.

IV.4.5 *Existe en $K[X]$ una infinidad de polinomios irreducibles, primos entre sí dos a dos.*

Demostración. Sean P_1, \dots, P_n polinomios irreducibles, no asociados dos a dos. Sea $A = P_1 P_2 \dots P_n + 1 = Q + 1$. Es claro que $(A, Q) = (1)$, luego ningún divisor de A está asociado a ninguno de los P_i ($1 \leq i \leq n$). Puesto que A posee un divisor irreducible P_{n+1} (según IV.4.3) se deduce que, si existen n polinomios irreducibles, no asociados, dos a dos, existen $n + 1$. Ahora bien, al menos, existe uno (los polinomios de grado 1 son irreducibles); luego existe una infinidad. c.q.d.

Nota importante. La propiedad de irreductibilidad depende del cuerpo K de los coeficientes (contrariamente a la propiedad de ser el mcd o el mcm).

Por ejemplo, $X^2 - 2$ es irreducible en $\mathbb{Q}[X]$ y

$$X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2})$$

en $\mathbb{R}[X]$.

§ IV.5 FUNCIÓN POLINOMIO. RAÍCES. FÓRMULA DE TAYLOR

● De momento, K designa a un anillo conmutativo. Más adelante, supondremos de nuevo que K es un cuerpo.

⁽¹⁾ La existencia de los α_P viene del hecho de que $P^0 = 1$ divide a A (luego el conjunto \mathcal{E} de los enteros k tales que P^k divide a A es no vacío) y de que si P^k divide a A , esto implica $k \times \text{gr}(P) \leq \text{gr}(A)$ (luego \mathcal{E} está acotado superiormente). Cf. teorema I.10.1.

DEFINICIÓN IV. 5.1

Sea $P = a_0 + a_1 X + \dots + a_n X^n$ un polinomio con coeficientes en K , y L una K -álgebra unífera, de elemento unidad I . Se le llama **función polinomio** ⁽¹⁾ asociado a P en L , y se designa por \tilde{P}_L , a la aplicación de L en L definida por:

$$\tilde{P}_L(x) = a_0 I + a_1 x + \dots + a_n x^n \quad (x \in L).$$

Si L' es una sub- K -álgebra de L , la restricción de \tilde{P}_L a L' es $\tilde{P}_{L'}$.

Introducimos la K -álgebra $\mathcal{F}(L)$ de las aplicaciones de L en L . Las propiedades de base de la aplicación $P \mapsto \tilde{P}_L$, de $K[X]$ en $\mathcal{F}(L)$ se resumen en el siguiente:

TEOREMA IV.5.1

Para toda K -álgebra L , la aplicación $P \mapsto \tilde{P}_L$ de $K[X]$ en $\mathcal{F}(L)$ es un **homomorfismo** de K -álgebras. En otras palabras, para todo par de polinomios $P, Q \in K[X]$ y todo $\lambda \in K$, se tiene:

$$(\widetilde{P + Q})_L = \tilde{P}_L + \tilde{Q}_L, \quad \widetilde{P_L Q_L} = \tilde{P}_L \tilde{Q}_L, \quad \widetilde{\lambda P}_L = \lambda \tilde{P}_L.$$

Demostración. Las leyes de composición en $K[X]$ se han definido precisamente para que las fórmulas anteriores fuesen verdaderas. c.q.d.

Casos particulares

1) Se toma $L = K$. La función \tilde{P}_K se designa simplemente por \tilde{P} , y se le llama *función polinomio asociada a P* . Este caso será el más importante en nuestro estudio.

2) Se toma $L = K[X]$. Para todo polinomio $Q \in K[X]$, el polinomio $\tilde{P}_L(Q)$ es, precisamente, el polinomio $P \circ Q$ (ver § 1). Para $Q = X$, $\tilde{P}_L(Q)$ es el mismo polinomio P . Por esta razón, un polinomio P con coeficientes en K , se puede designar indiferentemente por P o por $P(X)$.

3) Se toma $L = \mathcal{F}(K)$. Si I designa a la aplicación idéntica de K en K , $\tilde{P}_L(I)$ es precisamente la función \tilde{P} , que se vuelve a encontrar así «globalmente».

Nota. Si L designa una K -álgebra cualquiera, resulta de IV.5.1 que, para cada $u \in L$, la aplicación $P \mapsto \tilde{P}_L(u)$, de $K[X]$ en L , es un homomorfismo de K -

⁽¹⁾ Se denomina también, frecuentemente, **función polinómica**. (N. del. T.)

álgebras, y la imagen de este homomorfismo es la intersección de las sub- K -álgebras de L que contienen a u . A esta subálgebra de L se le llama *subálgebra engendrada por u* , y se le designa a menudo por $K[u]$ (ver Cap. XI).]

En lo que sigue de este párrafo, nos interesaremos únicamente por el homomorfismo $\tau : P \mapsto \tilde{P}$, de $K[X]$ en $\mathcal{F}(K)$.

En general, τ no es *inyectiva ni epiyectiva*. Dicho de otro modo, existen aplicaciones de K en K que no son funciones polinomios, por una parte; y por otra, la relación $\tilde{P}_1 = \tilde{P}_2$ no implica $P_1 = P_2$.

Ejemplos

1) La aplicación $f : x \mapsto \frac{1}{x^2 + 1}$ de \mathbf{R} en \mathbf{R} no es una función polinomio, pues $f \neq 0$, $\lim_{x \rightarrow +\infty} f(x) = 0$ y un polinomio $\neq 0$ no admite límite 0 cuando $x \rightarrow +\infty$.

2) Sea K un cuerpo finito y designemos por $\alpha_1, \alpha_2, \dots, \alpha_q$ a los elementos de K . Hagamos $P = 0$ (polinomio nulo en $K[X]$) y $Q = \prod_{k=1}^q (X - \alpha_k)$. Es claro que $\tilde{P} = 0$, y se tiene $\tilde{Q} = 0$ en virtud de IV.5.1. Sin embargo $Q \neq 0$, puesto que el grado de Q es q . Precizando, en este caso el núcleo del homomorfismo $P \mapsto \tilde{P}$, es el ideal $(Q) = (X^q - X)$ de $K[X]$ (cf. § 6, ejemplo 4).

Los polinomios P_1 y P_2 son *funcionalmente idénticos* si $\tilde{P}_1 = \tilde{P}_2$, *formalmente idénticos* si son iguales (en $K[X]$, es decir, si tienen los mismos coeficientes). El ejemplo (2) anterior demuestra que *la identidad funcional no implica, en general, la identidad formal* (si bien, es claro que el recíproco es verdadero). Se observará, sin embargo, que los polinomios de grado cero se identifican con las funciones constantes.

● Supondremos de nuevo que de ahora en adelante K es un cuerpo conmutativo.

Además, de ahora en adelante designaremos por $P(a)$ al valor de la función polinomio \tilde{P} asociada a P en el punto $a \in K$, en vez de designarlo por $\tilde{P}(a)$.

Una *raíz* de $P \in K[X]$ en K es un elemento $a \in K$ tal que $P(a) = 0$.

TEOREMA IV.5.2

- a) Si K es un cuerpo conmutativo, y $a \in K$ y $P \in K[X]$, para que P sea divisible por el polinomio $X - a$ es necesario y suficiente que a sea una raíz de P .
- b) Si $P \in K[X]$, $\text{gr}(P) \leq n$, y si P posee, en K , $n + 1$ raíces distintas, se tiene $P = 0$ formalmente.

Demostración

a) La división euclídea de P por $X - a$ da

$$P = (X - a) Q + R \quad (R = \text{Cte}).$$

En virtud de la proposición IV.5.1, se tiene

$$P(a) = (X - a)(a) \cdot Q(a) + R = R, \text{ pues } (X - a)(a) = 0.$$

Luego (P es divisible por $X - a$) $\Leftrightarrow (R = 0)$, de ahí el resultado.

b) Razonemos por recurrencia sobre n . La propiedad es evidente si $n = 0$, puesto que P es entonces una constante. Supongámosla verdadera para el entero $n - 1 \geq 0$, y demostrémosla para el entero n . A este fin, designemos por a_1, a_2, \dots, a_{n+1} , a $n + 1$ raíces distintas de P en K . Según la parte a) de la demostración, se puede escribir

$$P(X) = (X - a_{n+1}) P_1(X), \text{ en donde } P_1 \text{ designa un nuevo polinomio.}$$

Puesto que $\text{gr}(P) = \text{gr}(X - a_{n+1}) + \text{gr}(P_1)$, vemos que $\text{gr}(P_1) \leq n - 1$.

Substituimos X por a_i en la relación $P(X) = (X - a_{n+1}) P_1(X)$, para $1 \leq i \leq n$. Se obtiene: $(a_i - a_{n+1}) P_1(a_i) = 0$, de donde $P_1(a_i) = 0$ puesto que $a_i - a_{n+1} \neq 0$. En virtud de la hipótesis de recurrencia, se deduce que $P_1 = 0$, de donde $P = 0$. \square

Notas

1) La parte a) de la demostración es válida si se supone únicamente que K es un *anillo conmutativo unífero cualquiera*.

● 2) La parte b) de la demostración es válida si A es un *anillo conmutativo íntegro*; además era previsible teniendo en cuenta III.6.6 y el final del § IV.1.

COROLARIO

\parallel Si K es infinito, la identidad formal de polinomios es equivalente a la identidad funcional de estos polinomios.

En efecto, supongamos $P \neq 0$ y $\tilde{P} = 0$; si $n = \text{gr}(P)$, es posible encontrar $n + 1$ elementos distintos en K , que son raíces de P puesto que $\tilde{P} = 0$, lo que contradice IV.5.2. Luego $(\tilde{P} = 0)$ implica $(P = 0)$.

Mejor aún: el anterior razonamiento demuestra que *dos polinomios que toman el mismo valor para una infinidad de elementos de K son formalmente idénticos*. (Esta nota es la base de todos los razonamientos prácticos de identificación.)

Así cuando K es infinito, la aplicación $\tau : K[X] \rightarrow \mathcal{F}(K)$ es inyectiva, y es posible identificar, con la ayuda de τ , $K[X]$ con una subálgebra sobre K de $\mathcal{F}(K)$.

En otras palabras, no hay motivos para distinguir un polinomio de la función polinomio asociada.

Ejemplo

Sea $S_m = 2 \cos m\theta$ (m entero ≥ 1) y $x = S_1 = 2 \cos \theta$. Un cálculo elemental demuestra que, para $m \geq 3$, $S_m = xS_{m-1} - S_{m-2}$. La fórmula

$$(1) \quad S_m = x^m + \sum_{2 \leq 2q \leq m} (-1)^q \left[\binom{m-q}{q} + \binom{m-q-1}{q-1} \right] x^{m-2q}$$

se demuestra, entonces, por recurrencia.

La fórmula de De Moivre da directamente:

$$S_m = 2 \operatorname{Re} (\cos \theta + i \operatorname{sen} \theta)^m = 2 \sum_{0 \leq 2k \leq m} \binom{m}{2k} (-1)^k \operatorname{sen}^{2k} \theta \cos^{m-2k} \theta ;$$

$$\operatorname{sen}^{2k} \theta = (1 - \cos^2 \theta)^k$$

$$S_m = 2 \sum_{\substack{0 \leq \lambda \leq k \\ 0 \leq 2k \leq m}} (-1)^{k-\lambda} \binom{m}{2k} \binom{k}{\lambda} \cos^{m-2(k-\lambda)} \theta =$$

$$= 2 \sum_{\substack{0 \leq 2q \leq m \\ 2q \leq 2k \leq m}} (-1)^q \binom{m}{2k} \binom{k}{k-q} \cos^{m-2q} \theta$$

$$(2) \quad S_m = 2 \sum_{\substack{0 \leq 2q \leq m \\ 2q \leq 2k \leq m}} (-1)^q \binom{m}{2k} \binom{k}{k-q} \frac{1}{2^{m-2q}} x^{m-2q}$$

en (1) y (2) se ha remplazado $x = 2 \cos \theta$ por una indeterminada X . Obtenemos dos polinomios formales que toman el mismo valor para una infinidad de números reales, a saber los números de la forma $x = 2 \cos \theta$, es decir, los números del intervalo $[-2, 2]$. Luego estos polinomios son formalmente idénticos, lo que significa que en (1) y (2) se pueden «identificar» los coeficientes de las mismas potencias de x . De ahí las relaciones

$$\sum_{0 \leq 2k \leq m} \binom{m}{2k} = 2^{m-1},$$

y, para $2 \leq 2q \leq m$,

$$\sum_{2q \leq 2k \leq m} \binom{m}{2k} \binom{k}{k-q} \frac{1}{2^{m-2q+1}} = \binom{m-q}{q} + \binom{m-q-1}{q-1}.$$

* Una aplicación del teorema IV.5.2

Vamos a demostrar el resultado clásico siguiente: si K es un cuerpo conmutativo, todo subgrupo finito del grupo multiplicativo $K^* = K \setminus \{0\}$ es cíclico.

A este fin, designemos por G a uno de estos subgrupos, por n al orden de G , y por m al mayor entero igual al orden de uno de los elementos de G . Se trata de demostrar que $m = n$. Supongamos que $m < n$, luego m será un divisor estricto de n . Los elementos $x \in K$ cuyo orden divide a m son, según las propiedades del orden de un elemento, las raíces del polinomio $P = X^m - 1$; puesto que $\text{gr}(P) = m$, el número de estos elementos, según IV.5.2, a lo sumo m . De esto resulta que existe un elemento $b \in G$, en que el orden p no divide a m .

Utilizando la descomposición de m y p en factores primos, vemos que existe un número primo q , los enteros α, β que verifican $0 \leq \alpha < \beta$, y los enteros m', p' no divisibles por q , tales que:

$$m = q^\alpha m' \quad p = q^\beta p'.$$

Designemos por a a un elemento de G de orden m . Entonces a^{q^α} tiene orden m' y $b^{p'}$ tiene orden q^β (pues si x es un elemento de orden uv , x_u es de orden v). Los números m' y q^β son primos entre sí, de lo que se deduce (cf. Prop. II.4.3) que $a^{q^\alpha} b^{p'}$ es de orden $m' q^\beta > m$, lo cual contradice la definición de m . c.q.d.

En esta demostración, la conmutatividad de K ha desempeñado un papel esencial (utilización del T. IV.5.2, y cálculo del orden de $a^{q^\alpha} b^{p'}$). Es fácil de ver que la propiedad enunciada resulta falsa en el caso de un cuerpo no conmutativo (cf. problema n.º 5).

Derivada

Sea $P = a_0 + a_1 X + \dots + a_n X^n$ un polinomio. Por definición, la derivada formal de P (designada por $\frac{dP}{dX}$, P' , o $P'(X)$) es el polinomio

$$a_1 + 2 a_2 X + \dots + n a_n X^{n-1}, \quad \text{si } n \geq 1.$$

Y si $n = 0$ o $-\infty$, $P' = 0$. La aplicación $P \mapsto P'$ es lineal de $K[X]$ en sí mismo; finalmente se tiene la fórmula siguiente:

$$(3) \quad (PQ)' = P' Q + P Q'.$$

La derivada $P^{(k)}$ de orden cualquiera $k \in \mathbb{N}$ de P se define por recurrencia por $P^{(0)} = P$, $P^{(k+1)} = (P^{(k)})'$. La aplicación $P \mapsto P^{(k)}$ es lineal de $K[X]$ en sí mismo. Por recurrencia sobre k , se deduce de (3) la fórmula de Leibnitz:

$$(4) \quad (PQ)^{(k)} = \sum_{0 \leq j \leq k} \binom{k}{j} P^{(j)} Q^{(k-j)}.$$

— Pongamos $Y = X + h$, los $(X^n)_{n \geq 0}$ forman una base del K -espacio vectorial $K[X]$. Puesto que $X = Y - h$, los $(Y^n)_{n \geq 0}$ constituyen un sistema de generadores de este espacio. Este sistema es libre, pues de $\sum_{k=0}^m \lambda_k Y^k = 0$, se deduce

$$\sum_{k=0}^m \lambda_k (X + h)^k = 0,$$

de donde se obtienen las relaciones

$$\sum_{k=p}^m \lambda_k \binom{k}{p} h^{k-p} = 0, \quad (p = 0, 1, 2, \dots, m).$$

Para $p = m$, esta relación se reduce a $\lambda_m = 0$, y se demuestra por recurrencia que todos los λ_k son nulos.

En resumen, cualquiera que sea $h \in K$, los $(Y^n)_{n \geq 0}$ forman una base de $K[X]$. Siendo esto así, la fórmula de Taylor nos proporciona las coordenadas de P en la base $(Y^n)_{n \geq 0}$.

Multiplicidad de una raíz

Sea a una raíz del polinomio P ; $X - a$ es un factor irreducible de P , luego existe un entero $\alpha > 0$ tal que $(X - a)^\alpha$ divide a P y $(X - a)^{\alpha+1}$ no divide a P . O bien, si escribimos $P = (X - a)^\alpha Q$, $X - a$ no divide a Q , lo que equivale a $Q(a) \neq 0$.

DEFINICIÓN IV.5.2

$\left\{ \begin{array}{l} \text{Se conservan las anteriores notaciones. El entero } \alpha \text{ es la } \mathbf{multiplicidad} \\ \text{de la raíz } a \text{ (definición válida para un cuerpo conmutativo cualquiera).} \end{array} \right.$

Por convenio, si a no es una raíz de P , entonces la multiplicidad de a es 0. La fórmula de Taylor permite caracterizar la multiplicidad de una raíz.

TEOREMA IV.5.4

$\left\| \begin{array}{l} \text{Sea } a \text{ una raíz del polinomio } P. \text{ Para que el orden de multiplicidad de } a \\ \text{sea } k, \text{ es necesario y suficiente que se verifiquen las relaciones siguientes:} \\ (6) \quad P(a) = P'(a) = \dots = P^{(k-1)}(a) = 0, \\ (7) \quad P^{(k)}(a) \neq 0. \end{array} \right.$

Demostración. La fórmula de Taylor da

$$(8) \quad P(X) = P(a) + \frac{X-a}{1!} P'(a) + \cdots + \frac{(X-a)^{k-1}}{(k-1)!} P^{(k-1)}(a) + \\ + \frac{(X-a)^k}{k!} P^{(k)}(a) + (X-a)^{k+1} S,$$

en donde S es un polinomio tal que $\text{gr}(S) = \text{gr}(P) - k - 1$.

Haciendo

$$(X-a)S + \frac{P^{(k)}(a)}{k!} = Q, \quad \text{y} \quad R = P(a) + \frac{X-a}{1!} P'(a) + \cdots + \frac{(X-a)^{k-1}}{(k-1)!} P^{(k-1)}(a),$$

la anterior relación se convierte en

$$P = (X-a)^k Q + R,$$

que es la división euclídea de P por $(X-a)^k$.

Luego

$$((X-a)^k \text{ divide a } P) \Leftrightarrow (R=0) \Leftrightarrow (P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0)$$

(cf. las notas que siguen al teorema IV.5.3); (6) es, pues, el conjunto de condiciones necesarias y suficientes para que a sea de orden $\geq k$.

Si a es de orden $\geq k$, $(X-a)^{k+1}$ divide a P si, y sólo si, $P^{(k)}(a) = 0$, puesto que $Q(a) = \frac{1}{k!} P^{(k)}(a)$. De donde se sigue el teorema. c.q.d.

Si K es de característica $p > 0$, el teorema IV.5.4 es falso. Por ejemplo, sea $K = \mathbf{Z}/5\mathbf{Z}$ y $P(X) = X^{10} + \bar{1} = (X - \bar{2})^5 (X - \bar{3})^5$ (\bar{k} designa a la clase de k mód (5)). P admite dos raíces de orden 5, y por consiguiente $P'(X) = 10X^9 = 0$ (puesto que la característica es 5), luego para todo $k \geq 1$, $P^{(k)}(X) = 0$.

Sin embargo, en este caso, volviendo a la demostración de IV.5.3 se observa fácilmente que (8) es válido si $k < p$. El teorema IV.5.4 conserva su validez en característica p , siempre que $k < p$. En particular, se tendrá presente el siguiente resultado:

IV.5.5 Si K es un cuerpo conmutativo, para que $a \in K$ sea una raíz **simple** || del polinomio $P \in K[X]$, es necesario y suficiente que $P(a) = 0$ y $P'(a) \neq 0$.

§ IV.6 RELACIONES ENTRE LOS COEFICIENTES
Y LAS RAÍCES.
DESCOMPOSICIÓN EN $\mathbf{C}[X]$ Y $\mathbf{R}[X]$

Aquí K designa un cuerpo conmutativo cualquiera. Sea $P \in K[X]$ un polinomio de grado n tal que todos sus factores irreducibles sean de grado 1. Se puede escribir

$$(1) \quad P = a_n(X - x_1)(X - x_2) \dots (X - x_n),$$

en donde las $x_i \in K$ son las raíces de P , entendiéndose que si x_{i_0} es una raíz de orden k , se repetirá k veces el término x_{i_0} en la sucesión x_1, x_2, \dots, x_n . Introduzcamos los coeficientes (a_i) de P :

$$(2) \quad P = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0 \quad (a_n \neq 0)$$

y desarrollemos (1). Se obtienen las relaciones

$$(3) \quad \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} x_{i_1} x_{i_2} \dots x_{i_k} = (-1)^k \frac{a_{n-k}}{a_n} \quad (1 \leq k \leq n)$$

llamadas *relaciones entre los coeficientes y las raíces*. En particular, para $k = 1$, se obtiene $x_1 + x_2 + \dots + x_n = -\frac{a_{n-1}}{a_n}$; y, para $k = n$:

$$x_1 x_2 \dots x_n = (-1)^n \frac{a_0}{a_n}.$$

Ejemplos y aplicaciones

1) En $\mathbf{C}[X]$ sea $P(X) = (X + 1)^n - e^{2nia}$. Las raíces de P son

$$x_k = 2i e^{i\left(\frac{k\pi}{n} + a\right)} \operatorname{sen}\left(\frac{k\pi}{n} + a\right), \quad 0 \leq k \leq n-1.$$

Según (3) se tiene

$$S = x_0 x_1 \dots x_{n-1} = (-1)^n (1 - e^{2ina});$$

y directamente se encuentra

$$S = 2^n i^n \exp\left[i \sum_{k=0}^{n-1} \left(\frac{k\pi}{n} + a\right)\right] \prod_{k=0}^{n-1} \operatorname{sen}\left(a + \frac{k\pi}{n}\right).$$

Luego

$$\sum_{k=0}^{n-1} \left(\frac{k\pi}{n} + a \right) = na + \frac{\pi}{2} (n-1).$$

Si se escribe

$$T = \prod_{k=0}^{n-1} \operatorname{sen} \left(\frac{k\pi}{n} + a \right)$$

se tiene, entonces,

$$S = 2^n i^n e^{ina} e^{\frac{i\pi}{2}(n-1)} T = -2^n (-1)^n i e^{ina} T$$

de donde $T = \frac{\operatorname{sen} na}{2^{n-1}}$, identidad que se puede encontrar también por medio de las fórmulas de Euler.

2) Sea p un número primo ≥ 3 ,

$$K = \mathbf{Z}/p\mathbf{Z} = [\bar{0}, \bar{1}, \bar{2}, \dots, \overline{p-1}].$$

Se sabe (cf. ejemplo 2, § III.6, p. 116) que todo $x \in K$ verifica $x^p = x$; el polinomio $P = X^p - X$ admite como raíces a todos los elementos de K , lo mismo que $Q = X(X - \bar{1})(X - \bar{2}) \dots (X - \overline{p-1})$. Luego $P - Q$ es de grado $\leq p-1$. Según IV.5.2, se tiene $P - Q = 0$, de donde

$$X(X^{p-1} - \bar{1}) = X^p - X = X(X - \bar{1})(X - \bar{2}) \dots (X - \overline{p-1}).$$

Dado que $K[X]$ es íntegro, se deduce:

$$(4) \quad X^{p-1} - \bar{1} = (X - \bar{1})(X - \bar{2}) \dots (X - \overline{p-1}).$$

El producto de las raíces es $\overline{((p-1)!)}$, y puesto que p es impar, (3) da

$$\overline{(p-1)!} = -\bar{1},$$

y de ahí la congruencia entre enteros:

$$(p-1)! + 1 \equiv 0 \pmod{p}.$$

Esta condición (que es evidentemente suficiente) es también necesaria para que p sea primo ≥ 3 (teorema de Wilson).

3) A modo de ejercicio, busquemos si el polinomio $X^2 + 1$ posee una raíz en $K = \mathbf{Z}/p\mathbf{Z}$ (p , número primo ≥ 3).

K^* es un grupo multiplicativo con $p - 1$ elementos, y $\varphi_2 : u \rightarrow u^2$ es un endomorfismo de K^* . El núcleo N_2 de φ_2 está formado por los $u \in K^*$ tales que $u^2 - 1 = 0$, luego $N_2 = \{-1, +1\}$. De esto se sigue que la imagen de φ_2 tiene $\frac{p-1}{2}$ elementos. Se tiene (puesto que p es impar)

$$X^{p-1} - 1 = (X^{\frac{p-1}{2}} - 1)(X^{\frac{p-1}{2}} + 1),$$

luego, siempre según (4), cada polinomio $X^{\frac{p-1}{2}} - 1$ y $X^{\frac{p-1}{2}} + 1$ posee $\frac{p-1}{2}$ raíces en K^* . Sea $u_0 \in \varphi_2(K^*)$, existe v_0 tal que $u_0 = v_0^2$, de donde $u_0^{\frac{p-1}{2}} = v_0^{p-1} = 1$, luego todo elemento de $\varphi_2(K^*)$ es raíz de $X^{\frac{p-1}{2}} - 1$. Puesto que

$$\text{card}(\varphi_2(K^*)) = \frac{p-1}{2} = \text{gr}(X^{\frac{p-1}{2}} - 1)$$

vemos que los elementos de $\varphi_2(K^*)$ son exactamente las raíces de $X^{\frac{p-1}{2}} - 1$. En particular, para que $(-1) \in \varphi_2(K^*)$, es necesario y suficiente que $(-1)^{\frac{p-1}{2}} = 1$: es la condición necesaria y suficiente para que $X^2 + 1$ posea una raíz en $\mathbf{Z}/p\mathbf{Z}$ (criterio de Euler sobre el carácter cuadrático de $-1 \pmod{p}$).

4) Sea K un cuerpo finito con q elementos a_1, a_2, \dots, a_q . Probaremos que en $K[X]$, los polinomios $P = X^q - X$ y $Q = \prod_{i=1}^q (X - a_i)$ son iguales. Los elementos a_1, a_2, \dots, a_q son raíces de Q , y veamos que son asimismo raíces de P . El elemento nulo es, evidentemente, raíz de P . Si $a_i \neq 0$, se tiene: $a_i^{q-1} = 1$, pues el grupo multiplicativo de los elementos invertibles de K tiene $q - 1$ elementos (cf. corolario 2 de II.7.1), luego $a_i^q = a_i$, de donde se sigue nuestra afirmación.

Puesto que $P - Q$ es de grado $\leq q - 1$, vemos, como en el ejemplo 2, que $P - Q = 0$, de donde

$$X^q - X = \prod_{i=1}^q (X - a_i).$$

Busquemos ahora el núcleo del homomorfismo $P \mapsto \tilde{P}$ de $K[X]$ en $\mathcal{F}(K)$. Este núcleo es un ideal principal $\mathcal{O} = (N)$, y, puesto que la función polinomio asociada a $X^q - X$ es nula, $X^q - X$ es múltiplo de N . Sin embargo, IV.5.2 prueba que todo polinomio no nulo de \mathcal{O} debe poseer un grado $\geq q$. Se tiene, pues (salvo factores invertibles), $N = X^q - X$, e $\mathcal{O} = (X^q - X)$. Además, $P \mapsto \tilde{P}$ es epiyectiva (cf. ejercicio VII.11, *polinomio de interpolación de Lagrange*).

Descomposición en $C[X]$ y $R[X]$ **DEFINICIÓN IV.6.1**

$\left\{ \begin{array}{l} \text{Un cuerpo } K \text{ se llama } \textbf{algebraicamente cerrado} \text{ si todo polinomio no} \\ \text{constante } P \in K[X] \text{ posee, por lo menos, una raíz en } K. \end{array} \right.$

Esto equivale a decir que todo $P \in K[X]$ no constante se descompone en un producto de factores de grado 1:

$$P = a_n \prod_{k=1}^n (X - x_k) \quad (\text{los } x_k \text{ no son necesariamente distintos})$$

o que *todo polinomio irreducible de $K[X]$ es de grado 1.*

Si K es algebraicamente cerrado, K es infinito; en efecto, si K es finito

$$K = \{ a_1, \dots, a_q \},$$

el polinomio $1 + \prod_{k=1}^q (X - a_k)$ carece de raíces en K .

Ningún subcuerpo L de \mathbf{R} es algebraicamente cerrado, pues el polinomio $X^2 + 1$ no posee ninguna raíz en L .

Admitiremos el teorema que sigue, que se demostrará en el tomo II (Análisis).

TEOREMA IV.6.1

\parallel *El cuerpo \mathbf{C} de los números complejos es algebraicamente cerrado (D'Alembert).*

Su consecuencia es, que *todo polinomio no constante con coeficientes complejos se descompone en un producto de factores de primer grado.*

Nota. \mathbf{C} no es el menor cuerpo algebraicamente cerrado que contiene al cuerpo \mathbf{Q} . Al menor de estos cuerpos se le llama *clausura algebraica de \mathbf{Q}* . Es posible demostrar que la clausura algebraica de \mathbf{Q} está formada por el conjunto de los *números algebraicos*: por definición, un número complejo x es algebraico si existe un polinomio no nulo $P \in \mathbf{Q}[X]$ tal que $P(x) = 0$. Esto equivale a decir que existe un $P \in \mathbf{Z}[X]$ tal que $P \neq 0$ y $P(x) = 0$ (eliminar los denominadores). La clausura algebraica $\hat{\mathbf{Q}}$ de \mathbf{Q} es un conjunto *numerable*, es decir, existe una biyección de \mathbf{N} en $\hat{\mathbf{Q}}$ — mientras que esto no es verdadero para \mathbf{R} y, a fortiori, tampoco para \mathbf{C} (cf. tomo 2).

Polinomios con coeficientes reales

Consideremos ahora un polinomio $P \in \mathbf{R}[X]$. Si a es una raíz de P , su complejo conjugado \bar{a} es también una raíz de P .

Luego en $\mathbf{C}[X]$ se tiene:

$$(4) \quad P = a_n \prod_{k=1}^n (X - x_k) \quad (\text{los } x_k \text{ no son necesariamente distintos dos a dos})$$

de donde resulta que el número de raíces complejas no reales es par.

Sean x_1, x_2, \dots, x_r los elementos reales de la sucesión (x_k) ;

$\left. \begin{array}{l} x_{r+1}, \dots, x_{r+s} \\ x_{r+s+1} = \bar{x}_{r+1}, \dots, x_n = \bar{x}_{r+s} \end{array} \right\} \begin{array}{l} \text{los elementos no reales} \\ \text{de la sucesión } (x_k). \end{array}$

Se tiene

$$\begin{aligned} P &= a_n \prod_{k=1}^r (X - x_k) \prod_{l=1}^s (X - x_{r+l}) (X - \bar{x}_{r+l}) \\ &= a_n \prod_{k=1}^r (X - x_k) \prod_{l=1}^s (X^2 - 2\alpha_l X + \beta_l); \end{aligned}$$

con

$$\begin{aligned} 2\alpha_l &= x_{r+l} + \bar{x}_{r+l} \in \mathbf{R}, \\ \beta_l &= x_{r+l} \bar{x}_{r+l} = |x_{r+l}|^2 \in \mathbf{R}_+. \end{aligned}$$

Así el polinomio $X^2 - 2\alpha_l X + \beta_l$ tiene coeficientes reales y, puesto que es de grado 2 y carece de raíces reales, es irreducible en $\mathbf{R}[X]$ ⁽¹⁾. Se puede enunciar:

TEOREMA IV.6.2

|| *Los únicos polinomios irreducibles de $\mathbf{R}[X]$ son los de primer grado, y los de la forma $X^2 - 2\alpha X + \beta$ con $\alpha^2 - \beta < 0$.*

Ejemplos

Sea $P = X^4 + pX^2 + q$, $p, q \in \mathbf{R}$. Si $p^2 - 4q \geq 0$ el polinomio

$$U^2 + pU + q$$

posee dos raíces reales a_1, a_2 , de donde $P = (X^2 - a_1)(X^2 - a_2)$.

$X^2 - a_i$ es irreducible o no según que $a_i < 0$ o $a_i \geq 0$.

⁽¹⁾ Es preciso tener en cuenta el hecho de que un polinomio sin raíces, en general, no es irreducible. Por ejemplo, $(X^2 + 1)^2$ carece de raíces en \mathbf{R} pero no es irreducible en $\mathbf{Q}[X]$.

Si $p^2 - 4q < 0$, a fin de descomponer P en $\mathbf{R}[X]$ es cómodo escribir, observando que $q > 0$:

$$P = (X^2 + \sqrt{q})^2 + (p - 2\sqrt{q})X^2.$$

Pero $p^2 - 4q < 0$, de donde $2\sqrt{q} - p > 0$, y

$$P = (X^2 + \sqrt{q})^2 - (2\sqrt{q} - p)X^2 = (X^2 + \sqrt{2\sqrt{q} - p}X + \sqrt{q}) \times \\ \times (X^2 - \sqrt{2\sqrt{q} - p}X + \sqrt{q}),$$

y estos dos factores son irreducibles puesto que P carece de raíces reales.

De este modo se pueden demostrar, sin dificultad, las fórmulas

$$X^6 + 1 = (X^2)^3 + 1^3 = (X^2 + 1)(X^4 - X^2 + 1) = (X^2 + 1) \times \\ \times (X^2 + \sqrt{3}X + 1)(X^2 - \sqrt{3}X + 1),$$

$$X^8 + 1 = (X^4 + 1)^2 - 2X^4 = (X^4 - \sqrt{2}X^2 + 1)(X^4 + \sqrt{2}X^2 + 1) = \\ P(X)P(-X),$$

con
$$P(X) = (X^2 + \sqrt{2 + \sqrt{2}}X + 1)(X^2 + \sqrt{2 - \sqrt{2}}X + 1).$$

Como ejercicio, el lector puede descomponer $X^{2^n} + 1$.

* § IV.7 NOCIONES SOBRE $K[X_1, X_2, \dots, X_n]$ (Polinomios con n variables o indeterminadas)

● Sea K un anillo conmutativo y n un entero ≥ 1 . Consideramos el conjunto \mathbf{N}^n de las sucesiones finitas $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ de n enteros ≥ 0 ; para $\alpha \in \mathbf{N}^n$ establecemos $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$, y para $\alpha, \beta \in \mathbf{N}^n$,

$$\alpha + \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n).$$

DEFINICIÓN IV.7.1

Un polinomio con n variables y con coeficientes en K es una familia $\{ (a_\alpha)_{\alpha \in \mathbf{N}^n} \}$ en la cual todos los términos son nulos, salvo un número finito. (a_α es, pues nulo si $\sup (\alpha_1, \alpha_2, \dots, \alpha_n)$ es suficientemente grande.)

Al conjunto de todos estos polinomios se le designa por $K[X_1, X_2, \dots, X_n]$. Se le dota de las siguientes leyes:

— *Adición en $K[X_1, \dots, X_n]$*

$$(a_\alpha) + (b_\alpha) = (a_\alpha + b_\alpha) \quad (\alpha \in \mathbf{N}^n).$$

Para esta ley, $K[X_1, X_2, \dots, X_n]$ es un *grupo abeliano*. El elemento neutro es $0 = (a_\alpha)_{\alpha \in \mathbf{N}^n}$ tal que $\forall \alpha, a_\alpha = 0$ (*polinomio nulo*). El opuesto de $P = (a_\alpha)$ es $-P = (-a_\alpha)$.

— *Multiplicación en $K[X_1, \dots, X_n]$*

Sean $P = (a_\alpha)$, $Q = (b_\beta)$ dos polinomios. La sucesión $R = (c_\gamma)_{\gamma \in \mathbf{N}^n}$ tal que

$$c_\gamma = \sum_{\alpha + \beta = \gamma} a_\alpha b_\beta$$

es un polinomio. Existen, en efecto, dos enteros n_0, n_1 tales que $a_\alpha = 0$ para $|\alpha| > n_0$ y $b_\beta = 0$ para $|\beta| > n_1$. Para $|\gamma| > n_0 + n_1$ se tiene entonces $c_\gamma = 0$ (en efecto, puesto que $|\gamma| = |\alpha| + |\beta|$, cada término $(a_\alpha b_\beta)$ tiene entonces un factor nulo por lo menos, luego es nulo).

Se escribe $R = P \cdot Q$.

El producto es asociativo, conmutativo y distributivo respecto de la adición: estas propiedades se deducen de las propiedades análogas del anillo K (cf. § 1).

El elemento (a_α) tal que $a_\alpha = 1$ para $\alpha = (0, 0, \dots, 0)$, y $a_\alpha = 0$ para $\alpha \neq (0, 0, \dots, 0)$ es neutro para el producto. Se le designará por 1, pues no se presta a confusión.

Resumiendo, se puede enunciar:

TEOREMA IV.7.1

|| Dotado de las leyes anteriores, el conjunto $K[X_1, \dots, X_n]$ es un anillo conmutativo unífero.

A este anillo también se le designa por $K[X_1, \dots, X_n]$.

Inclusión de K en $K[X_1, \dots, X_n]$

Sea $j : K \mapsto K[X_1, \dots, X_n]$ tal que

$$j(\lambda) = (a_\alpha) : \begin{aligned} a_\alpha &= \lambda & \text{si } \alpha &= (0, \dots, 0), \\ a_\alpha &= 0 & \text{si } \alpha &\neq (0, 0, \dots, 0). \end{aligned}$$

j es un homomorfismo de anillos, inyectivo, por medio del cual se identifica K con un cierto subanillo de $K[X_1, \dots, X_n]$ (cf. § 1). Dotamos a $K[X_1, \dots, X_n]$ de la ley «externa»

$$\begin{aligned} K \times K[X_1, \dots, X_n] &\rightarrow K[X_1, \dots, X_n] \\ (\lambda, P) &\mapsto \lambda P = j(\lambda) \cdot P. \end{aligned}$$

Entonces el anillo $K[X_1, \dots, X_n]$ se convierte en una K -álgebra.

Para cada $i = 1, 2, \dots, n$, definimos el polinomio X_i (variable) por medio de la sucesión $(a_\alpha)_{\alpha \in \mathbb{N}^n}$:

$$\begin{aligned} a_\alpha &= 1 \quad \text{si } \alpha = (\alpha_1, \dots, \alpha_n) \text{ satisface a } \alpha_i = 1, \text{ y } \alpha_j = 0 \text{ para } j \neq i; \\ a_\alpha &= 0 \quad \text{en los otros casos.} \end{aligned}$$

Con estas notaciones el polinomio $P = (a_\alpha)_{\alpha \in \mathbb{N}^n}$ se escribe:

$$(1) \quad P = \sum_{\alpha \in \mathbb{N}^n} a_\alpha X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n};$$

(la suma tiene sentido pues sólo existe un número finito de términos $a_\alpha \neq 0$) y la expresión (1) es única ya que los coeficientes son precisamente los a_α . Se dice que los polinomios $(X_i)_{1 \leq i \leq n}$ son generadores de la K -álgebra $K[X_1, \dots, X_n]$.

En la expresión (1) ordenamos los términos respecto de las potencias crecientes de X_n ; se obtiene

$$(2) \quad P = \sum_{k \in J} A_k X_n^k \quad \text{en donde } J \text{ es una parte finita de } \mathbb{N}, \\ \text{y } A_k \text{ es un polinomio de } K[X_1, \dots, X_{n-1}].$$

Esta operación define a una inyección

$$\varphi : K[X_1, \dots, X_n] \rightarrow K[X_1, \dots, X_{n-1}][X_n]$$

en el anillo de los polinomios con una variable, con coeficientes en $K[X_1, \dots, X_{n-1}]$; φ es un homomorfismo de anillos (por la propia definición de suma y de producto en $K[X_1, \dots, X_n]$), y evidentemente φ es inyectiva, luego biyectiva. De donde:

TEOREMA IV.7.2

|| La ordenación según las potencias crecientes de X_n determina un isomorfismo canónico de $K[X_1, \dots, X_n]$ en $K[X_1, \dots, X_{n-1}][X_n]$.

El teorema IV.7.2 permite utilizar los teoremas de «permanencia» para demostrar propiedades de $K[X_1, \dots, X_n]$. Por ejemplo:

COROLARIO

|| Si K es íntegro, $K[X_1, \dots, X_n]$ es íntegro.

En efecto, hemos visto ya que si K es íntegro, $K[X]$ es íntegro. Es suficiente, pues, para probar el corolario, razonar por recurrencia sobre n utilizando el teorema IV.4.2.

Grado parcial

Sea k un entero comprendido entre 1 y n . El grado de $P \in K[X_1, \dots, X_n]$ respecto de X_k es el grado de P , considerado como polinomio en X_k , con coeficientes en $K[X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n]$. Si designamos por $\text{gr}_k(P)$ dicho grado, en virtud del § 1, tendremos:

$$(3) \quad \text{gr}_k(P + Q) \leq \sup(\text{gr}_k(P), \text{gr}_k(Q)),$$

$$(4) \quad \text{gr}_k(PQ) \leq \text{gr}_k(P) + \text{gr}_k(Q).$$

Si K es íntegro, en (4) se tiene la igualdad.

El grado parcial en X_k es también la mayor potencia con la que X_k interviene en un monomio $a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}$ de P , si $P \neq 0$. (Si $P = 0$, se recuerda que $\text{gr}_k P = -\infty$.)

Grado total

Sea $P = \sum_{\alpha \in \mathbf{N}^n} a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}$. El grado total $\text{gr}(P)$ de P es $-\infty$ si $P = 0$, y $\sup_{a_\alpha \neq 0} |\alpha|$ si $P \neq 0$. Se tienen las fórmulas

$$(5) \quad \text{gr}(P + Q) \leq \sup(\text{gr}(P), \text{gr}(Q)),$$

$$(6) \quad \text{gr}(PQ) \leq \text{gr}(P) + \text{gr}(Q).$$

Si K es íntegro, en (6) se tiene la igualdad.

Demostración. Hagamos

$$P = \sum_{\alpha \in \mathbf{N}^n} a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n},$$

$$Q = \sum_{\alpha \in \mathbf{N}^n} b_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}, \quad p = \text{gr}(P), \quad q = \text{gr}(Q).$$

Por definición, se tiene $a_\alpha = 0$ para $|\alpha| > p$ y $b_\alpha = 0$ para $|\alpha| > q$, de donde $a_\alpha + b_\alpha = 0$ para $|\alpha| > \sup(p, q)$, lo que prueba (5).

Deduciremos (6) utilizando una propiedad de los *polinomios homogéneos* que se establecerá más adelante, pero cuya demostración es independiente.

Agrupando, en P y Q , los monomios del mismo grado total, es posible escribir P y Q de forma única, de la manera siguiente:

$$P = P_0 + P_1 + \cdots + P_p, \quad Q = Q_0 + Q_1 + \cdots + Q_q,$$

en donde, para todo $k = 0, 1, \dots, p$ [resp. para todo $j = 0, 1, \dots, q$], P_k es homogéneo de grado k [resp. Q_j es homogéneo de grado j]. Se tiene, pues:

$$PQ = \sum_{k=0}^p \sum_{j=0}^q P_k Q_j.$$

Luego, según las propiedades de los polinomios homogéneos, el producto $P_k Q_j$ es nulo u homogéneo de grado $k + j$. Se tiene, pues,

$$\text{gr}(P_k + Q_j) \leq k + j \leq p + q,$$

de donde (6) se obtiene por aplicación de (5) a la suma $\sum_{k,j} P_k Q_j$.

Cuando K es íntegro, sabemos que $K[X_1, X_2, \dots, X_n]$ es íntegro (corolario de IV.7.2). En este caso, con las notaciones que preceden, se tiene: $P_p Q_q \neq 0$, de donde $\text{gr}(P_p Q_q) = p + q$; y $P_p Q_q$ representa la suma de los monomios de grado total $p + q$ en PQ . Se tiene, en consecuencia,

$$\text{gr}(PQ) = p + q = \text{gr}(P) + \text{gr}(Q). \text{ c.q.d.}$$

Derivadas parciales

- Sea $P \in K[X_1, \dots, X_n]$; la derivada parcial $\frac{\partial P}{\partial X_k}$ (o P'_{X_k}) es la derivada de P respecto de X_k , cuando se considera a P como un polinomio en X_k con coeficientes en $K[X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n]$.
- La aplicación $\frac{\partial}{\partial X_k}: K[X_1, \dots, X_n] \rightarrow K[X_1, \dots, X_n]$ es lineal y verifica

$$\frac{\partial}{\partial X_k}(PQ) = \frac{\partial P}{\partial X_k} \cdot Q + P \frac{\partial Q}{\partial X_k}.$$

IV.7.3 Cualesquiera que sean $i, j = 1, 2, \dots, n$, se tiene:

$$\left\| \begin{array}{l} (7) \end{array} \right. \quad \frac{\partial}{\partial X_i} \circ \frac{\partial}{\partial X_j} = \frac{\partial}{\partial X_j} \circ \frac{\partial}{\partial X_i}$$

Por la linealidad, es suficiente comprobar (7) para monomios; pero en este caso (7) es evidente. \square

De (7) resulta que se pueden componer, en cualquier orden, varias derivaciones. Por definición, la aplicación

$$\frac{\partial^\alpha}{\partial X_1^{\alpha_1} \dots \partial X_n^{\alpha_n}} : K[X_1, \dots, X_n] \rightarrow K[X_1, \dots, X_n], \quad \alpha = (\alpha_1, \dots, \alpha_n),$$

es la aplicación compuesta que se obtiene al componer α_1 veces consigo misma la $\frac{\partial}{\partial X_1}$, con α_2 consigo misma la $\frac{\partial}{\partial X_2}$, etc...

Es, pues, K -lineal, y si $P = \sum_{\alpha \in \mathbb{N}^n} a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}$, se tiene

$$(8) \quad \frac{\partial^\alpha P}{\partial X_1^{\alpha_1} \dots \partial X_n^{\alpha_n}}(0) = \alpha_1! \alpha_2! \dots \alpha_n! a_\alpha.$$

Polinomios homogéneos

Sea $P = \sum_{\alpha \in \mathbb{N}^n} a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}$ un polinomio. P es *homogéneo de grado k* si la relación $a_\alpha \neq 0$ implica la relación $|\alpha| = k$. En otras palabras, todos los monomios de P tienen grado total k .

Equivale a decir que, en la K -álgebra $K[X_1, X_2, \dots, X_n, Y]$ (y $K[X_1, \dots, X_n]$ se halla identificado con una sub- K -álgebra de $K[X_1, \dots, X_n, Y]$), se tiene la relación formal:

$$(9) \quad P(YX_1, YX_2, \dots, YX_n) = Y^k P(X_1, \dots, X_n).$$

El conjunto formado por 0 y por los polinomios homogéneos de grado k forma un sub- K -módulo de $K[X_1, \dots, X_n]$. Se obtiene inmediatamente:

Si P es homogéneo de grado k , y Q es homogéneo de grado l , PQ es homogéneo de grado $k + l$, o nulo.

Si P es homogéneo de grado k , se verifica inmediatamente la fórmula

$$(10) \quad X_1 P'_{X_1} + X_2 P'_{X_2} + \dots + X_n P'_{X_n} = kP \quad (\text{fórmula de Euler}).$$

En general, el recíproco es falso. Sin embargo, *es verdadero si K es un cuerpo de característica nula* (pues, en este caso la relación $P'_{X_k} = 0$ implica que P es independiente de X_k : la misma demostración que en el curso de Análisis).

Fórmula de Taylor

- Aquí supondremos que K es un cuerpo de característica nula.

Notación. Si $P \in K[X_1, \dots, X_n]$, la p -ésima potencia simbólica de P es el polinomio con $2n$ variables $X_1, \dots, X_n, U_1, \dots, U_n$:

$$\sum_{|\alpha|=p} \frac{\alpha!}{\alpha_1! \dots \alpha_n!} \cdot \frac{\partial^\alpha P}{\partial X_1^{\alpha_1} \dots \partial X_n^{\alpha_n}} U_1^{\alpha_1} \dots U_n^{\alpha_n}$$

que se le designa por $[U_1 P'_{X_1} + \dots + U_n P'_{X_n}]^{(p)}$. Si $p = 0$, por convenio, la p -ésima potencia simbólica de P es P . Con estas notaciones, se tiene:

TEOREMA IV.7.4

$$\left\| \begin{array}{l} \text{Para todo polinomio } P \in K[X_1, \dots, X_n], \text{ de grado total } \leq m, \text{ se tiene} \\ P(X_1 + U_1, X_2 + U_2, \dots, X_n + U_n) = \sum_{k=0}^m \frac{1}{k!} [U_1 P'_{X_1} + \dots + U_n P'_{X_n}]^{(k)} \\ \text{(fórmula de Taylor).} \end{array} \right.$$

Demostración. En virtud de la linealidad de las derivadas, es suficiente comprobar esta fórmula en el caso en que P es un monomio.

$$X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}, \quad \alpha_1 + \alpha_2 + \dots + \alpha_n = p \leq m.$$

Pero en este caso, la fórmula resulta de la fórmula del binomio, aplicada a cada factor $(X_i + U_i)^{\alpha_i}$ del producto

$$S = (X_1 + U_1)^{\alpha_1} (X_2 + U_2)^{\alpha_2} \dots (X_n + U_n)^{\alpha_n},$$

y de la relación (8):

$$S = \prod_{i=1}^n \left(\sum_{\beta_i + \gamma_i = \alpha_i} \frac{\alpha_i!}{\beta_i! \gamma_i!} X_i^{\beta_i} U_i^{\gamma_i} \right) = \sum_{\substack{\beta_1 + \gamma_1 = \alpha_1 \\ \vdots \\ \beta_n + \gamma_n = \alpha_n}} \frac{\alpha_1! \alpha_2! \dots \alpha_n!}{\beta_1! \beta_2! \dots \beta_n! \gamma_1! \dots \gamma_n!} \times \\ \times X_1^{\beta_1} \dots X_n^{\beta_n} U_1^{\gamma_1} \dots U_n^{\gamma_n}.$$

Dejamos los detalles al lector.]]

Función polinomio

Sea $P = \sum_{\alpha \in \mathbb{N}^n} a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}$ un polinomio con coeficientes en el anillo K . Se le asocia la función polinomio $\tilde{P}: K^n \rightarrow K$ tal que $\tilde{P}(x_1, x_2, \dots, x_n)$ sea el elemento obtenido reemplazando X_i por x_i en la expresión de P .

La aplicación

$$\begin{aligned} \Phi: K[X_1, \dots, X_n] &\rightarrow \mathcal{F}(K^n, K) \\ P &\mapsto \tilde{P} \end{aligned}$$

(en donde $\mathcal{F}(K^n, K)$ designa al álgebra de las aplicaciones de K^n en K) es un homomorfismo de álgebras, y se tiene el siguiente teorema, análogo al corolario del teorema IV.5.2:

TEOREMA IV.7.5

Si K es un cuerpo infinito, el homomorfismo Φ definido anteriormente es inyectivo (dicho en otras palabras, la identidad funcional entre polinomios con n variables equivale a su identidad formal).

Demostración. Por recurrencia sobre n . Para $n = 1$, es el teorema IV.3.1. Supongamos el teorema verdadero para $n - 1$, y sea

$$P \in K[X_1, \dots, X_n] = K[X_1, \dots, X_{n-1}][X_n].$$

Ordenemos P según las potencias crecientes de X_n ; se obtiene:

$$P = \sum_{k=0}^m A_k X_n^k, \text{ en donde } A_k \in K[X_1, \dots, X_{n-1}].$$

Supongamos $P \neq 0$. Existe k_0 tal que $A_{k_0} \neq 0$. Según la hipótesis de recurrencia existe $(x_1, x_2, \dots, x_{n-1}) \in K^{n-1}$ tal que $\tilde{A}_{k_0}(x_1, \dots, x_{n-1}) \neq 0$.

El polinomio con *una* variable.

$$Q(X_n) = \sum_{k=0}^m \tilde{A}_k(x_1, \dots, x_{n-1}) X_n^k$$

no es formalmente nulo. Luego existe $x_n \in K$ tal que $\tilde{Q}(x_n) \neq 0$, es decir,

$$\tilde{P}(x_1, \dots, x_n) \neq 0. \text{ Así } P \neq 0 \text{ implica } \tilde{P} \neq 0. \text{ c.q.d.}$$

Nota. De hecho, el anterior razonamiento prueba un resultado más fuerte que IV.7.5: si E es una parte infinita de K , un polinomio $P \in K[X_1, \dots, X_n]$, nulo para todas las n -plas (x_1, \dots, x_n) de elementos de E , es formalmente nulo.

Por ejemplo, si P no es formalmente nulo en el anillo $\mathbf{Q}[X_1, \dots, X_n]$, existen enteros m_1, m_2, \dots, m_n tales que $P(m_1, m_2, \dots, m_n) \neq 0$.

Factorización en $K[X_1, \dots, X_n]$, $n \geq 2$ (K cuerpo conmutativo)

Es fácil de ver que $K[X_1, \dots, X_n]$ no es un anillo principal. Por ejemplo, sea \mathcal{O} el ideal de $K[X, Y]$ engendrado por X e Y : $\mathcal{O} = (X, Y)$. Si \mathcal{O} fuera principal, X e Y serían múltiplos de un polinomio P tal que $\mathcal{O} = (P)$. Los únicos polinomios que dividen a la vez a X y a Y son las constantes. Se tendría, pues, que $P = 1$, lo cual es absurdo, ya que \mathcal{O} está formado por todos los polinomios sin término constante.

No se pueden, pues, desarrollar las propiedades aritméticas de

$$K[X_1, \dots, X_n],$$

para $n \geq 2$, tal como se ha hecho para $n = 1$. Así el teorema de Bezout no se puede aplicar; en $K[X, Y]$ los polinomios X e Y sólo poseen a las constantes como factores comunes, y, por lo tanto, cualesquiera que sean $A, B \in K[X, Y]$, $AX + BY$ es no constante.

Sin embargo, las propiedades aritméticas esenciales son verdaderas para

$$K[X_1, \dots, X_n]$$

en particular el *teorema de Gauss*, y la *descomposición en factores irreducibles*. Con mayor precisión, demostraremos más adelante (Cap. XIV) el siguiente:

TEOREMA IV.7.6

Todo polinomio P con n variables, con coeficientes en el cuerpo K , admite una factorización de la forma

$$P = P_1^{\alpha_1} P_2^{\alpha_2} \dots P_k^{\alpha_k},$$

en donde $(\forall i) \alpha_i \in \mathbf{N}^$, y en donde los P_i son polinomios dos a dos irreducibles sin factores comunes.*

Además, dos descomposiciones de P de este tipo difieren sólo en el orden de los factores, y en un factor constante.

Polinomios invariantes frente a un grupo de permutaciones

Sea G un subgrupo del grupo \mathfrak{S}_n de las biyecciones del conjunto \mathbf{N}_n^* en sí mismo. Diremos que el polinomio $P \in K[X_1, \dots, X_n]$ es invariante respecto de G si, para toda permutación $\sigma \in G$, se tiene:

$$P(X_1, \dots, X_n) = P(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(n)}).$$

Es claro que si H es un subgrupo de G , todo polinomio invariante para G es invariante para H .

El conjunto de todos los polinomios invariante para G forma un subanillo \mathcal{P}_G de $K[X_1, \dots, X_n]$; si H es un subgrupo de G , $\mathcal{P}_G \subset \mathcal{P}_H$.

$\mathcal{P}_{\{e\}} = K[X_1, \dots, X_n]$ (e , permutación idéntica).

● $\mathcal{P}_{\mathfrak{S}_n}$ es el anillo de los polinomios simétricos: a sus elementos se les llama **polinomios simétricos**.

Ejemplo de polinomios simétricos:

$$S_k = X_1^k + X_2^k + \dots + X_n^k \quad (k \in \mathbf{N})$$

$$\sigma_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} X_{i_1} X_{i_2} \dots X_{i_k} \quad (1 \leq k \leq n).$$

(A los σ_k se les llama polinomios simétricos elementales.)

● Sea \mathcal{A}_n el grupo alternado de orden n , formado por las permutaciones pares de \mathfrak{S}_n , entonces $\mathcal{P}_{\mathcal{A}_n}$ es el anillo de los polinomios alternados, y sus elementos son los **polinomios alternados**.

Ejemplo: el polinomio $\prod_{1 \leq i < j \leq n} (X_j - X_i)$ es alternado y no simétrico (cf. demostración II.7.3) siempre que K sea de característica $\neq 2$.

● Sea Γ un grupo de permutaciones engendrado por un ciclo de longitud n : \mathcal{P}_Γ es un anillo de polinomios cíclicos.

Ejemplo: el polinomio $X_1 X_2^2 + X_2 X_3^2 + \dots + X_{n-1} X_n^2 + X_n X_1^2$ es cíclico (para la permutación circular $\begin{pmatrix} 1 & 2 & \dots & n \\ 2 & 3 & \dots & 1 \end{pmatrix}$).

Como ejercicio vamos a caracterizar todos los polinomios homogéneos de grado k y alternados, suponiendo que K es de característica $\neq 2$.

Sea P un polinomio de este tipo, y para todo $\sigma \in \mathfrak{S}_n$ designemos por P_σ al polinomio $P(X_{\sigma(1)}, \dots, X_{\sigma(n)})$; si $\sigma \in \mathcal{A}_n$, $P_\sigma = P$; si σ es impar y τ es par, $P_{\sigma\tau} = P_\sigma$. Si P no es simétrico, y si σ es impar, el polinomio P_σ no depende de σ ; designando por Q este polinomio, se tiene que $Q \neq P$.

El polinomio $R = P - Q$ es alternado, homogéneo de grado k . Si σ es una permutación impar, se tiene: $R_\sigma = -R$. Nos vemos, pues, obligados a estudiar un

polinomio alternado R , que se transforme en su opuesto por medio de las permutaciones impares.

Si $i < j$, se tiene, pues (ya que τ_{ij}^n es impar):

$$R(X_1, \dots, X_i, \dots, X_j, \dots, X_n) = -R(X_1, \dots, X_j, \dots, X_i, \dots, X_n).$$

En esta relación, reemplazamos X_i y X_j por la variable U . Dado que K no tiene característica 2, resulta:

$$(11) \quad R(X_1, \dots, U, \dots, U, \dots, X_n) = 0.$$

Esto implica que R es divisible por $X_j - X_i$. En efecto, R se puede escribir:

$$(12) \quad R = \sum_{\alpha + \beta = k} A_{\alpha\beta} X_i^\alpha X_j^\beta;$$

y puesto que $A_{\alpha\beta}$ son polinomios respecto a las $X_l (l \neq i, j)$ (11) significa exactamente que $\sum_{\alpha + \beta = k} A_{\alpha\beta} = 0$. Luego (12) se puede también escribir (restando $\sum A_{\alpha\beta} X_i^k$):

$$R = \sum_{\alpha + \beta = k} A_{\alpha\beta} X_i^\alpha (X_j^\beta - X_i^\beta).$$

Entonces nuestra afirmación resulta evidente, puesto que

$$X_j^\beta - X_i^\beta = (X_j - X_i) \left(\sum_{\lambda=0}^{\beta-1} X_j^{\beta-1-\lambda} X_i^\lambda \right)$$

es divisible por $X_j - X_i$.

R es, pues, divisible por $X_j - X_i$ para $i < j$. Dado que los $X_j - X_i$ son primos entre sí dos a dos, si admitimos el teorema XIV.1.3 (que generaliza al IV.2.6 en $K[X_1, \dots, X_n]$), se deduce que R es divisible por $A = \prod_{1 \leq i < j \leq n} (X_j - X_i)$, o sea

$$R = AR_1.$$

Tomemos entonces $\sigma \in \mathfrak{S}_n$. Si $\sigma \in \mathcal{A}_n$, σ deja invariantes a R y a A , luego también deja invariante a R_1 . Si σ es impar, cambia a R y a A en sus opuestos, luego también deja invariante a R_1 ; en otras palabras, R_1 es simétrico.

Recapitulando, tenemos: $P - Q = AR_1$. Pero, por otro lado, $R_0 = P + Q$ es evidentemente simétrico. Las dos relaciones

$$R_0 = P + Q, \quad AR_1 = P - Q, \quad \text{dan:} \quad P = \frac{1}{2}(R_0 + AR_1),$$

de donde, cambiando las notaciones, se sigue el resultado:

Todo polinomio alternado es de la forma $S + AT$, en donde S y T son polinomios simétricos, y en donde $A = \prod_{1 \leq i < j \leq n} (X_j - X_i)$.

Nota. Este último resultado es falso en característica 2. Por ejemplo, si el cuerpo K de base es $\mathbf{Z}/2\mathbf{Z}$, en $K[X, Y, T]$ el polinomio $XY^2 + YT^2 + TX^2$ es alternado pero no es de la forma $S + AT$ antes indicada.

Capítulo V

Funciones simétricas. Ecuaciones algebraicas (teoría elemental)

El viejo problema consistente en resolver ecuaciones algebraicas con la ayuda de fórmulas en que interviniesen únicamente funciones racionales de «radicales» fue esclarecido por Galois (1811-1832). La respuesta es, en general, negativa para las ecuaciones de grado ≥ 5 . En este capítulo se desarrollan los métodos de resolución algebraica de las ecuaciones de grado 3 y 4, cuyo interés es, sobre todo, teórico e histórico: el intento por resolver estas ecuaciones condujo, paso a paso, a partir del siglo xv, a la introducción de los números complejos (cf. § 2).

No debe olvidarse que la búsqueda de los *valores numéricos aproximados* de las raíces de una ecuación (incluso de grado 3 ó 4) es mucho más rápida con la ayuda de técnicas clásicas de análisis numérico (método de Newton, etc.) que mediante el cálculo teórico expuesto a continuación (que conduce a pesadas operaciones numéricas).

Por el contrario, el estudio de las funciones simétricas de las raíces, por el que comenzaremos, presenta un gran interés, tanto para la teoría de Galois como para el estudio clásico de las curvas.

● En este capítulo se consideran polinomios con coeficientes en un cuerpo K algebraicamente cerrado ⁽¹⁾, de característica 0.

§ V.1 POLINOMIOS SIMÉTRICOS

Hemos visto, al final del capítulo IV, que los polinomios

$$\sigma_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} X_{i_1} X_{i_2} \dots X_{i_k} \quad (1 \leq k \leq n)$$

⁽¹⁾ Esta hipótesis acerca de K no quita generalidad pues, según el teorema de Steinitz, todo cuerpo conmutativo se puede sumergir como subcuerpo en un cuerpo algebraicamente cerrado.

de $K[X_1, \dots, X_n]$ son simétricos. En lo que sigue probaremos que todo polinomio simétrico puede expresarse por medio de los σ_k .

TEOREMA V.1.1

Sea F un polinomio con n variables. Definimos el polinomio G con n variables por medio de

$$(1) \quad G(X_1, X_2, \dots, X_n) = F(\sigma_1(X), \sigma_2(X), \dots, \sigma_n(X)),$$

en donde $\sigma_k(X) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} X_{i_1} X_{i_2} \dots X_{i_k}$.

Si G es nulo, F es nulo.

Demostración. Para cada $x = (x_1, \dots, x_n) \in K^n$, establecemos

$$\tilde{\sigma}(x) = (\tilde{\sigma}_1(x), \tilde{\sigma}_2(x), \dots, \tilde{\sigma}_n(x)),$$

en donde $\tilde{\sigma}_k(x)$ designa el valor del polinomio σ_k en x . Entonces la aplicación

$$\tilde{\sigma} : K^n \rightarrow K^n$$

es epiyectiva. En efecto, sea $y = (y_1, \dots, y_n) \in K^n$. La ecuación

$$X^n - y_1 X^{n-1} + y_2 X^{n-2} - \dots + (-1)^n y_n = 0$$

posee n raíces (distintas o no), a saber x_1, \dots, x_n .

Hagamos $x = (x_1, x_2, \dots, x_n)$, y es claro que $\tilde{\sigma}(x) = y$ (relaciones entre los coeficientes y las raíces). Pero este resultado, junto con (1), prueba que la función polinomio F es nula (ya que $G = 0$), luego que F es formalmente nulo (puesto que al ser nula su característica, K es infinito, cf. p. 108). \square

COROLARIO

Sea G un polinomio con n variables. Si existe un polinomio

$$F \in K[X_1, \dots, X_n]$$

tal que $G(X_1, \dots, X_n) = F(\sigma_1(X), \sigma_2(X), \dots, \sigma_n(X))$, F es único.

Designemos por p el grado total del polinomio simétrico G . Se puede escribir,

de forma única,

$$G(X_1, \dots, X_n) = \sum_{k=0}^p G_k(X_1, \dots, X_n),$$

en donde G_k es un polinomio homogéneo de grado total k . Evidentemente cada G_k es simétrico, pues una permutación sobre los X_i no cambia el grado total de un monomio. Para expresar G con la ayuda de los polinomios σ_i , será, pues, suficiente expresar los G_k . Dicho de otra manera, podemos suponer que G es homogéneo de grado total p , y simétrico. El grado parcial de G respecto de X_i es entonces independiente de i , $1 \leq i \leq n$.

DEFINICIÓN V.1.1

Sea G un polinomio simétrico con n variables, homogéneo de grado total p . Al entero p se le llama **peso** de G . Al grado de G respecto de una cualquiera de las variables X_i se le llama **orden** de G ; se designa por $\omega(G)$.

TEOREMA V.1.2

Designamos por G un polinomio de $K[X_1, \dots, X_n]$, simétrico y homogéneo de peso $p \geq 1$.
a) Existe un polinomio $\Phi \in K[X_1, \dots, X_n]$, único, tal que

$$G(X_1, \dots, X_n) = \Phi(\sigma_1(X), \sigma_2(X), \dots, \sigma_n(X));$$

b) el grado total de Φ es igual al orden de G ;
c) todo monomio no nulo $\lambda \sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \dots \sigma_n^{\alpha_n}$ de Φ es tal que

$$(1) \quad \alpha_1 + 2\alpha_2 + 3\alpha_3 + \dots + n\alpha_n = p.$$

Demostración. La unicidad de Φ resulta del teorema V.1.1. Todo consiste, pues, en establecer la existencia y las propiedades de Φ . Para ello, razonamos por recurrencia sobre el entero $\delta(G) = p + n$. Para $p = 1$ el teorema es verdadero para todo n , puesto que entonces G es de la forma $\lambda \sigma_1$. Para $n = 1$ el teorema es verdadero para todo p , pues $\sigma_1 = X_1$.

Supongamos, pues, cierto el teorema para todos los polinomios simétricos homogéneos H tales que $\delta(H) \leq \delta(G) - 1$, y supongamos además

$$n \geq 2, p \geq 2.$$

El polinomio $G_1(X_1, \dots, X_{n-1}) = G(X_1, \dots, X_{n-1}, 0)$ es simétrico respecto de las variables X_1, \dots, X_{n-1} (en efecto, las permutaciones de \mathbf{N}_{n-1}^* se identifican con las permutaciones de \mathbf{N}_n^* que dejan fijo el entero n).

1) Supongamos primeramente que $G_1 = 0$, y entonces X_n es un factor de G . Puesto que G es simétrico, cada X_i es también un factor y G es divisible por el producto de los X_i ; se tiene pues:

$$G = \sigma_n(X) \cdot G_2(X_1, \dots, X_n);$$

y el polinomio G_2 es simétrico de peso $p - n$, y $\delta(G_2) = p$. Según la hipótesis de recurrencia, existe $\Phi_2 \in K[X_1, \dots, X_n]$ tal que

$$G_2(X_1, \dots, X_n) = \Phi_2(\sigma_1(X), \sigma_2(X), \dots, \sigma_n(X)).$$

El grado total de Φ_2 es igual al orden de G_2 , o sea $\omega(G_2)$, y:

$$\omega(G) = \omega(G_2) + 1.$$

El polinomio $\Phi \in K[U_1, U_2, \dots, U_n]$ definido por

$$\Phi(U_1, \dots, U_n) = U_n \Phi_2(U_1, \dots, U_n)$$

es de grado total $\omega(G_2) + 1 = \omega(G)$. Puesto que se tiene

$$G(X_1, \dots, X_n) = \Phi(\sigma_1(X), \dots, \sigma_n(X)),$$

las propiedades a) y b) del teorema quedan demostradas en este caso.

La propiedad c) se verifica fácilmente: según la hipótesis de recurrencia, todo monomio $\lambda \sigma_1^{\alpha_1} \dots \sigma_n^{\alpha_n}$ de Φ_2 es tal que $\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = p - n$. Este monomio origina el monomio $\lambda \sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \dots \sigma_{n-1}^{\alpha_{n-1}} \sigma_n^{1+\alpha_n}$, y se tiene entonces:

$$\alpha_1 + 2\alpha_2 + \dots + (n-1)\alpha_{n-1} + n(\alpha_n + 1) = p.$$

2) Supongamos $G_1 \neq 0$. Entonces el grado total de G_1 es igual a p , de donde

$$\delta(G_1) = p + n - 1,$$

y $\omega(G_1) = \omega(G)$. Sean $\sigma'_1, \sigma'_2, \dots, \sigma'_{n-1}$ los polinomios simétricos elementales de X_1, \dots, X_{n-1} , que se obtienen substituyendo X_n por 0 en

$$\sigma_1, \sigma_2, \dots, \sigma_{n-1}.$$

Según la hipótesis de recurrencia, existe un polinomio Φ_1 , de grado total $\omega(G)$, tal que

$$\begin{aligned} G_1(X_1, \dots, X_{n-1}) &= \\ &= \Phi_1(\sigma'_1(X_1, \dots, X_{n-1}), \sigma'_2(X_1, \dots, X_{n-1}), \dots, \sigma'_{n-1}(X_1, \dots, X_{n-1})); \end{aligned}$$

el polinomio simétrico con n variables

$$G_2(X_1, \dots, X_n) = G(X_1, \dots, X_n) - \Phi_1(\sigma_1(X), \dots, \sigma_{n-1}(X))$$

es o bien nulo, o bien de peso p (pues $\sigma_1, \dots, \sigma_{n-1}$ tienen el mismo grado total que $\sigma'_1, \dots, \sigma'_{n-1}$). Si $G_2 = 0$, el teorema está demostrado (pues la propiedad (c) se verifica inmediatamente). Si $G_2 \neq 0$, el polinomio G_2 se anula cuando se substituye X_n por 0. En virtud de la parte 1) de la demostración, existe

$$\Phi_2 \in K[X_1, \dots, X_n],$$

de grado total $\omega(G) - 1$, tal que $G_2(X_1, \dots, X_n) = \sigma_n(X) \Phi_2(\sigma_1(X), \dots, \sigma_n(X))$.

En este último caso, se tiene entonces

$$G(X_1, \dots, X_n) = \Phi_1(\sigma_1(X), \dots, \sigma_{n-1}(X)) + \sigma_n(X) \Phi_2(\sigma_1(X), \dots, \sigma_n(X)),$$

en donde Φ_1 es de grado total $\omega(G)$, y Φ_2 de grado total $\omega(G) - 1$. Dado que U_n no figura en Φ_1 , vemos que el polinomio

$$\Phi(U_1, \dots, U_n) = \Phi_1(U_1, \dots, U_{n-1}) + U_n \Phi_2(U_1, \dots, U_n)$$

es de grado total $\omega(G)$ respecto de U_1, \dots, U_n . Finalmente, la propiedad c) del teorema es verdadera para los polinomios Φ_1 y Φ_2 (cuando $G_2 \neq 0$); el polinomio Φ la verifica sin dificultad en todos los casos. c.q.d.

Nota. Resulta también de la demostración V.1.2 que si G tiene todos sus coeficientes en un subanillo A de K , el polinomio Φ tiene igualmente todos sus coeficientes en A . (Por ejemplo, si G tiene coeficientes enteros, Φ también tiene coeficientes enteros.)

Nota. La demostración precedente no da el método práctico para hallar el polinomio Φ . Sin embargo, la unicidad de Φ , y las propiedades b) y c) del teorema V.1.2, permiten aligerar los cálculos en la búsqueda práctica de Φ .

Antes de dar ejemplos, precisemos algunas notaciones: Sean m un entero $\leq n$, $\alpha_1, \dots, \alpha_m$ enteros > 0 , α la sucesión finita $(\alpha_1, \dots, \alpha_m)$. \mathcal{F} designa el conjunto de las aplicaciones inyectivas $p \mapsto i_p$ de \mathbf{N}_m^* en \mathbf{N}_n^* .

● Por convenio, el polinomio $P_\alpha = \sum_{\tau \in \mathcal{F}} X_{\tau(1)}^{\alpha_1} \dots X_{\tau(m)}^{\alpha_m}$ se escribirá:

$$P_\alpha = \sum_{i_1, i_2, \dots, i_m} X_{i_1}^{\alpha_1} \dots X_{i_m}^{\alpha_m}; P_\alpha \text{ es de peso: } |\alpha| = \sum \alpha_i \text{ y de orden: } \sup_i (\alpha_i).$$

El número de términos de P_α es $\text{card}(\mathcal{F}) = \frac{n!}{(n-m)!}$.

Por ejemplo, para $n = 5$, los polinomios

$$\sum_{i_1, i_2, i_3} X_{i_1} X_{i_2} X_{i_3} \text{ y } \sum_{i_1, i_2, i_3} X_{i_1} X_{i_2} X_{i_3}^2$$

contienen ambos 60 términos.

● Designamos por k_1, \dots, k_p los valores distintos que toman los enteros a_1, \dots, a_m ($k_1 < k_2 < \dots < k_p$); las imágenes recíprocas J_i , por la aplicación $i \mapsto a_i$, de los elementos k_i , constituyen una partición de \mathbf{N}_m^* . Pongamos $v_i = \text{card}(J_i)$, $1 \leq i \leq p$. Puesto que v_i de los exponentes a_i son iguales a k_i , $v_i!$ términos de $P_a = \sum_{i_1, i_2, \dots, i_m} X_{i_1}^{a_1} \dots X_{i_m}^{a_m}$ son iguales a un término dado. Por consiguiente, todos los coeficientes del polinomio

$$\frac{1}{v_1! v_2! \dots v_p!} \sum_{i_1, i_2, \dots, i_m} X_{i_1}^{a_1} \dots X_{i_m}^{a_m}$$

son iguales a 1.

Por convenio, este polinomio, igual a $\frac{1}{v_1! v_2! \dots v_p!} P_a$, se designará por:

$$(2) \quad \sum X_1^{a_1} X_2^{a_2} \dots X_m^{a_m}.$$

Para $n = 5$, por ejemplo, el polinomio $\sum X_1 X_2 X_3$ contiene 10 términos, y el polinomio $\sum X_1 X_2 X_3^2$ contiene 30.

Obsérvese que $\sum_{i_1, \dots, i_m} X_{i_1}^{a_1} X_{i_2}^{a_2} \dots X_{i_m}^{a_m} = \sum X_1^{a_1} \dots X_m^{a_m}$; si, y sólo si, los a_i son distintos, dos a dos.

Ejemplos

1) Cálculo de $S = \sum X_i^3$ ($n \geq 3$).

Haremos intervenir los σ_k bajando el orden paso a paso; se tiene:

$$\sigma_1(\sum X_i^2) = S + \sum_{i,j} X_i^2 X_j \quad \text{además} \quad \sum X_i^2 = \sigma_1^2 - 2\sigma_2,$$

y para calcular $T = \sum_{i,j} X_i^2 X_j$, se continúa bajando el orden:

$$(\sum X_i)(\sum_{j,k} X_j X_k) = 2 \sum_{i,j} X_i^2 X_j + \sum_{i,j,k} X_i X_j X_k = 2T + 6\sigma_3,$$

de donde

$$T = \sigma_1 \sigma_2 - 3\sigma_3.$$

Lo cual nos da
$$S = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3.$$

2) Cálculo de $S = \sum X_1^2 X_2^2$ ($n \geq 4$).

Aquí procederemos *por identificación*, buscando S en forma de un polinomio en los σ_k .

S es de orden 2 y de peso 4. Luego la expresión de S sólo puede contener, según el teorema V.1.2, términos en $\sigma_1, \sigma_3, \sigma_2^2$ y σ_4 , únicos monomios en σ_i de grado ≤ 2 y de peso 4. De donde:

$$S = A\sigma_1\sigma_3 + B\sigma_2^2 + C\sigma_4.$$

Para la ecuación $X^4 + X^2 = 0$, se tiene $\sigma_1 = \sigma_3 = \sigma_4 = 0, \sigma_2 = 1, S = 1$; de donde

$B = 1$. Para la ecuación $X\left(X^3 - \frac{3}{2}X^2 + \frac{1}{2}\right) = 0$, se tiene

$$\sigma_1 = \frac{3}{2}, \quad \sigma_2 = \sigma_4 = 0, \quad \sigma_3 = -\frac{1}{2},$$

de donde $S = -\frac{3}{4}A$. Pero como las raíces de este polinomio son $0, 1, 1, -\frac{1}{2}$,

$S = \frac{3}{2}$; de donde $A = -2$. Puesto que no hay términos de orden 1 en S , se tiene

$A + C = 0$, de donde $C = 2$ y finalmente

$$S = -2\sigma_1\sigma_3 + \sigma_2^2 + 2\sigma_4$$

Se podrá comprobar esta fórmula para la ecuación $X(X-1)^3 = 0$.

3) Cálculo de $S = \sum X_1^3 X_2^2$ ($n \geq 4$).

Bajamos el orden paso a paso:

$$\left(\sum_i X_i^2\right)\left(\sum_{j,k} X_j X_k\right) = 2S + \sum_{i,j,k} X_i^2 X_j X_k = 2S + T,$$

$$6\sigma_1\sigma_3 = \left(\sum_i X_i\right)\left(\sum_{j,k,l} X_j X_k X_l\right) = 3T + \sum_{i,j,k,l} X_i X_j X_k X_l = 3T + 24\sigma_4;$$

de donde $T = 2\sigma_1\sigma_3 - 8\sigma_4, \quad 2S = (\sigma_1^2 - 2\sigma_2)2\sigma_2 - T$

$$S = \sigma_1^2\sigma_2 - 2\sigma_2^2 - \sigma_1\sigma_3 + 4\sigma_4.$$

4) Cálculo de $S = \sum X_1^2 X_2^2 X_3^2$ ($n \geq 6$).

Bajando el orden se tiene:

$$36 \sigma_3^2 = \left(\sum_{i_1, j_1, k_1} X_{i_1} X_{j_1} X_{k_1} \right) \left(\sum_{(i_2, j_2, k_2)} X_{i_2} X_{j_2} X_{k_2} \right) = 36 S +$$

$$+ 18 \sum_{i, j, k, l} X_i^2 X_j^2 X_k X_l + 9 \sum_{i, j, k, l, m} X_i^2 X_j X_k X_l X_m$$

$$+ \sum_{i, j, k, l, m, n} X_i X_j X_k X_l X_m X_n.$$

La relación precedente nos da

$$\sigma_3^2 = S + 2 T + 6 U + 20 \sigma_6, \quad \text{con } T = \sum X_1^2 X_2^2 X_3 X_4$$

y
$$U = \sum X_1^2 X_2 X_3 X_4 X_5.$$

Para calcular T y U , formamos sucesivamente:

$$a) \left(\sum_{i, j} X_i X_j \right) \left(\sum_{k, l, m, n} X_k X_l X_m X_n \right) = 48 \sigma_2 \sigma_4 = 12 \sum_{i, j, k, l} X_i^2 X_j^2 X_k X_l +$$

$$+ 8 \sum_{i, j, k, l, m} X_i^2 X_j X_k X_l X_m + 6! \sigma_6 = 48 T + (48 \times 4) U + 6! \sigma_6,$$

de donde

$$(3) \quad \sigma_2 \sigma_4 = T + 4 U + 15 \sigma_6;$$

$$b) \left(\sum X_i \right) \left(\sum_{j, k, l, m, n} X_j X_k X_l X_m X_n \right) = 5! \sigma_1 \sigma_5 = 5 \sum_{i, j, k, l, m} X_i^2 X_j X_k X_l X_m +$$

$$+ 6! \sigma_6 = 5! U + 6! \sigma_6,$$

luego

$$(4) \quad \sigma_1 \sigma_5 = U + 6 \sigma_6.$$

$$(3) \text{ y } (4) \text{ dan } U = \sigma_1 \sigma_5 - 6 \sigma_6, \quad T = \sigma_2 \sigma_4 - 4 \sigma_1 \sigma_5 + 9 \sigma_6.$$

La relación hallada al principio nos permite calcular S :

$$S = \sigma_3^2 - 2 \sigma_2 \sigma_4 + 2 \sigma_1 \sigma_5 - 2 \sigma_6.$$

Los procedimientos empleados en los ejemplos anteriores dan buenos resultados cuando es bajo el orden de la función simétrica estudiada.

§ V.2 FÓRMULAS DE NEWTON

Estas fórmulas permiten expresar de manera recurrente, los polinomios simétricos $S_k = \sum_{1 \leq p \leq n} X_p^k$ ($1 \leq k \leq n$) en función de los σ_k , o los σ_k en función de los S_k .

En particular, todo polinomio simétrico se puede escribir (de forma única) como un polinomio respecto de los S_k . Esta observación nos será útil, en los cálculos prácticos, cuando sea grande el orden de la función simétrica estudiada.

TEOREMA V.2.1

Los polinomios simétricos $S_k = \sum_{1 \leq p \leq n} X_p^k$, con n variables, verifican las siguientes relaciones (fórmulas de Newton):

a) para $k \geq n$ $S_k - \sigma_1 S_{k-1} + \sigma_2 S_{k-2} + \cdots + (-1)^n \sigma_n S_{k-n} = 0$,

b) para $k \leq n$ $S_k - \sigma_1 S_{k-1} + \sigma_2 S_{k-2} + \cdots + (-1)^k k \sigma_k = 0$,

con el convenio $S_0 = n$.

Demostración

a) Sea U una variable distinta de las X_i y consideremos el polinomio con $n + 1$ variables

$$P(U) = (U - X_1)(U - X_2) \cdots (U - X_n) = U^n - \sigma_1 U^{n-1} + \sigma_2 U^{n-2} + \cdots + (-1)^n \sigma_n.$$

Se tiene: $P(X_i) = 0$ ($1 \leq i \leq n$), y la fórmula a) se obtiene multiplicando la relación $P(X_i) = 0$ por X_i^{k-n} y sumando miembro a miembro las relaciones obtenidas.

b) Partimos de la relación

$$(5) \quad \frac{\partial P}{\partial U} = \sum_{k=1}^n \frac{P(U)}{U - X_k}$$

(que resulta de la regla de derivación del producto, aplicada a

$$P(U) = \prod_{k=1}^n (U - X_k).$$

Por una parte, se tiene:

$$(6) \quad \frac{\partial P}{\partial U} = nU^{n-1} - (n-1)\sigma_1 U^{n-2} + \cdots + (-1)^{n-1} \sigma_{n-1},$$

en virtud de la definición de derivada.

Por otra parte, vamos a calcular $\frac{P(U)}{U - X_k}$, utilizando la relación $P(X_k) = 0$:

$$P(U) = P(U) - P(X_k) = U^n - X_k^n + \cdots + (-1)^p \sigma_p (U^{n-p} - X_k^{n-p}) + \cdots + (-1)^{n-1} \sigma_{n-1} (U - X_k).$$

Puesto que $U^{n-p} - X_k^{n-p} = (U - X_k) \sum_{j=0}^{n-p-1} U^{n-p-1-j} X_k^j$, se deduce:

$$(7) \quad \frac{P(U)}{U - X_k} = U^{n-1} + \sum_{p=1}^{n-1} A_{p,k} U^{n-1-p}, \quad \text{con} \quad A_{p,k} = X_k^p + \sum_{j=1}^p (-1)^j X_k^{p-j} \sigma_j.$$

Sumando miembro a miembro las relaciones (7), llevándolo a (5) y teniendo en cuenta (6) se obtienen las fórmulas b) identificando las potencias de U . \square

Ejemplos

- 1) Evidentemente se tiene $S_1 = \sigma_1$, $S_2 = \sigma_1^2 - 2\sigma_2$.
Para calcular S_3 ($n \geq 3$) escribimos las fórmulas b):

$$S_2 - \sigma_1 S_1 + 2\sigma_2 = 0, \quad S_3 - \sigma_1 S_2 + \sigma_2 S_1 - 3\sigma_3 = 0,$$

se obtiene

$$S_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3.$$

Para calcular S_4 , es suficiente escribir la fórmula de Newton siguiente ($n \geq 4$):

$$S_4 - \sigma_1 S_3 + \sigma_2 S_2 - \sigma_3 S_1 + 4\sigma_4 = 0.$$

Resulta, teniendo en cuenta los resultados precedentes,

$$S_4 = \sigma_1^4 - 4\sigma_1^2\sigma_2 + 4\sigma_3\sigma_1 + 2\sigma_2^2 - 4\sigma_4.$$

- 2) Para calcular un polinomio simétrico en función de los σ_k , se puede *aumentar el orden* y expresarlo así con la ayuda de los S_k . Después las fórmulas de Newton permiten terminar el cálculo.

Por ejemplo, para $S = \sum_{i,j} X_i^3 X_j$ ($n \geq 4$), se tiene:

$$(\sum X_i^3)(\sum X_j) = \sum X_i^4 + S, \quad S = \sigma_1 S_3 - S_4,$$

y, teniendo en cuenta los valores de S_3 y de S_4 ,

$$S = \sigma_1^2\sigma_2 - \sigma_1\sigma_3 - 2\sigma_2^2 + 4\sigma_4.$$

Igualmente, el polinomio $S = \sum X_i^3 X_j^2 X_k$ se obtiene por los siguientes cálculos:

$$\begin{aligned}(\sum X_i^3 X_j^2)(\sum X_k) &= S + \sum X_i^4 X_j^2 + \sum X_i^3 X_j^3, \\(\sum X_i^4)(\sum X_j^2) &= S_4 S_2 = \sum X_i^6 + \sum X_i^4 X_j^2, \\(\sum X_i^3)(\sum X_j^3) &= S_3^2 = \sum X_i^6 + \sum X_i^3 X_j^3, \\(\sum X_i^3)(\sum X_j^2) &= S_3 S_2 = \sum X_i^3 X_j^2 + S_5.\end{aligned}$$

$$S = S_1 S_2 S_3 - S_1 S_5 - S_2 S_4 + 2 S_6 - S_3^2.$$

El cálculo de S_6 es bastante pesado. El lector comprobará que se obtiene:

$$S = -12 \sigma_6 + 7 \sigma_1 \sigma_5 + 4 \sigma_2 \sigma_4 - 3 \sigma_1^2 \sigma_4 - 3 \sigma_3^2 + \sigma_1 \sigma_2 \sigma_3$$

3) Cuando se pide calcular los *valores* de funciones simétricas para las raíces de una ecuación particular, los cálculos se simplifican frecuentemente. Por ejemplo, tenemos que calcular $S_7 = x_1^7 + x_2^7 + x_3^7$ para las raíces x_i de la ecuación: $x^3 + px + q = 0$. Aquí no es adecuado calcular la expresión general de S_7 y substituir en ella $\sigma_1, \sigma_2, \sigma_3$ por $0, p, -q$.

Es más rápido buscar el resto de X^7 módulo $(X^3 + pX + q)$, pues este resto toma para las x_i el mismo valor que X^7 . Se obtiene

$$X^7 = 2pqX^2 + (q^2 - p^3)X - p^2q \pmod{(X^3 + pX + q)},$$

de donde
$$S_7 = 2pqS_2 + (q^2 - p^3)S_1 - 3p^2q;$$

y puesto que $S_2 = \sigma_1^2 - 2\sigma_2 = -2p$, se tiene:

$$S_7 = -7p^2q.$$

Aplicación de los polinomios simétricos a la transformación de ecuaciones (cf. Cap. VI)

Por definición, la transformada de la ecuación $P(x) = 0$, en donde

$$P(X) = X^n + p_1 X^{n-1} + \dots + p_n = (X - x_1) \dots (X - x_n),$$

por la función racional $y = F(x)$, es la ecuación:

$$(Y - F(x_1))(Y - F(x_2)) \dots (Y - F(x_n)) = 0.$$

Más adelante veremos que, en teoría, el problema se puede reducir al caso en que F es un polinomio.

En este caso, al desarrollar la ecuación transformada, los coeficientes de las potencias de Y se obtienen como *funciones simétricas de las x_i* . En principio es posible calcular estos coeficientes con la ayuda de p_1, p_2, \dots, p_n .

Ejemplos

1) Transformar $x^3 + px^2 + qx + r = 0$ por $y = x + \frac{1}{x}$.

El desarrollo de $\prod_{i=1}^3 \left(y - \left(x_i + \frac{1}{x_i} \right) \right)$ conduce a la transformada

$$y^3 + Py^2 + Qy + R = 0,$$

con:

$$P = - \sum \left(x_1 + \frac{1}{x_1} \right),$$

$$Q = \sum x_1 x_2 + \sum \frac{1}{x_1 x_2} + \sum \left(\frac{x_1}{x_2} + \frac{x_2}{x_1} \right),$$

$$R = x_1 x_2 x_3 + \frac{1}{x_1 x_2 x_3} + \sum \frac{x_1 x_2}{x_3} + \sum \frac{x_1}{x_2 x_3}.$$

La ecuación cuyas raíces son las $\frac{1}{x_i}$ es $rx^3 + qx^2 + px + 1 = 0$; designando por s_1, s_2, s_3 a las funciones simétricas de las $\frac{1}{x_i}$, se tiene:

$$P = -\sigma_1 - s_1 = p + \frac{q}{r}$$

$$Q = \sigma_2 + s_2 + \sigma_1 s_1 - 3 = q + \frac{p}{r} + \frac{pq}{r} - 3$$

$$\left(\text{pues } \sum \left(\frac{x_1}{x_2} + \frac{x_2}{x_1} \right) + 3 = (x_1 + x_2 + x_3) \left(\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} \right) \right)$$

$$R = \sigma_3 + s_3 + \sigma_2 s_1 - 2\sigma_1 + s_2 \sigma_1 - 2s_1,$$

$$\left(\text{pues } \left(\sum x_1 x_2 \right) \left(\sum \frac{1}{x_1} \right) = \sum \frac{x_1 x_2}{x_3} + 2 \sum x_1 \right)$$

$$R = -r - \frac{1}{r} - \frac{q^2}{r} + 2p - \frac{p^2}{r} + 2\frac{q}{r}.$$

2) Sean x_1, x_2, x_3 las raíces de $x^3 + px^2 + qx + r = 0$. Hallar la ecuación cuyas raíces son $x_2^2 + x_3^2, x_3^2 + x_1^2, x_1^2 + x_2^2$.

Puesto que $x_1^2 + x_2^2 + x_3^2 = \sigma_1^2 - 2\sigma_2 = p^2 - 2q$, el problema consiste en transformar la ecuación por $y = p^2 - 2q - x^2$.

La ecuación cuyas raíces son los números $p^2 - 2q - y_i$ es la transformada de la ecuación dada por $z = x^2$. Es suficiente, pues, hacer $z = p^2 - 2q - y$ (para obtener la transformada por y) en la ecuación en z . En el capítulo VI, veremos un método general para transformar una ecuación por $y = x^k$. Acabamos los cálculos de nuestro ejemplo con la ayuda de las funciones simétricas.

$$x_1^2 + x_2^2 + x_3^2 = p^2 - 2q,$$

$$\sum x_1^2 x_2^2 = -2\sigma_1 \sigma_3 + \sigma_2^2 = -2pr + q^2,$$

$$x_1^2 x_2^2 x_3^2 = \sigma_3^2 = r^2.$$

La ecuación en z es, pues,

$$z^3 - (p^2 - 2q)z^2 + (q^2 - 2pr)z - r^2 = 0;$$

y volviendo a $y = p^2 - 2q - z$, se obtiene la ecuación

$$y^3 + Py^2 + Qy + R = 0,$$

con

$$P = -2(p^2 - 2q), \quad Q = p^4 - 4p^2q + 5q^2 - 2pr,$$

$$R = -p^2q^2 + 2p^3r - 4pqr + 2q^3 + r^2.$$

* Fórmulas de Waring

Las fórmulas de Newton no dan una expresión general de los σ_k en función de los S_k , o viceversa. Una expresión de este tipo se puede obtener, sin embargo, con la ayuda de la teoría de las series formales (cf. § VII.4).

Consideremos de nuevo la identidad

$$P(U) = (U - X_1)(U - X_2) \dots (U - X_n) = U^n - \sigma_1 U^{n-1} + \dots + (-1)^n \sigma_n.$$

Si Y designa una nueva variable, se deduce la identidad

$$(8) \quad (1 - X_1 Y)(1 - X_2 Y) \dots (1 - X_n Y) = 1 - \sigma_1 Y + \sigma_2 Y^2 - \dots + (-1)^n \sigma_n Y^n.$$

En la serie formal

$$(9) \quad \text{Log}(1 + X) = X - \frac{X^2}{2} + \dots + (-1)^{m-1} \frac{X^m}{m} + \dots,$$

substituimos $1 + X$ por el primer miembro de (8), y se obtiene por una parte (teniendo en cuenta las propiedades formales del Log):

$$\begin{aligned}\text{Log} [(1 - X_1 Y) \dots (1 - X_n Y)] &= \sum_{k=1}^n \text{Log} (1 - X_k Y) \\ &= \sum_{k=1}^n \left(- \sum_{m \geq 1} \frac{1}{m} X_k^m Y^m \right) = - \sum_{m \geq 1} \frac{1}{m} S_m Y^m\end{aligned}$$

$$\text{con } S_m = \sum_{k=1}^n X_k^m.$$

En (9), reemplazamos $1 + X$ por el segundo miembro de (8), y se obtiene, por adición, la relación formal:

$$\begin{aligned}- \text{Log} (1 - \sigma_1 Y + \sigma_2 Y^2 + \dots + (-1)^n \sigma_n Y^n) &= \\ &= \sum_{r \geq 1} \frac{1}{r} (\sigma_1 Y - \sigma_2 Y^2 + \dots + (-1)^n \sigma_n Y^n)^r.\end{aligned}$$

Poniendo $\sigma_i = (-1)^i p_i$, se obtiene, igualando las dos expresiones halladas:

$$\begin{aligned}(10) \quad - \text{Log} (1 + p_1 Y + \dots + p_n Y^n) &= \sum_{r \geq 1} \frac{(-1)^r}{r} (p_1 Y + p_2 Y^2 + \dots + p_n Y^n)^r \\ &= \sum_{m \geq 1} \frac{S_m}{m} Y^m.\end{aligned}$$

Aplicando a $(p_1 Y + p_2 Y^2 + \dots + p_n Y^n)^r$ la fórmula general del binomio e identificando los términos en Y^m en la relación anterior resulta

$$(11) \quad \boxed{\frac{S_m}{m} = \sum_{r_1 + 2r_2 + 3r_3 + \dots + nr_n = m} \frac{(-1)^{r_1 + r_2 + \dots + r_n}}{(r_1 + r_2 + \dots + r_n)} \frac{(r_1 + r_2 + \dots + r_n)!}{r_1! r_2! \dots r_n!} p_1^{r_1} p_2^{r_2} \dots p_n^{r_n}}.$$

Como la serie formal recíproca de $\text{Log} (1 + X)$ es $e^X - 1$, (10) nos da también

$$1 + p_1 Y + p_2 Y^2 + \dots + p_n Y^n = \exp \left(- \sum_{m \geq 1} \frac{S_m}{m} Y^m \right) = e^{-S_1 Y} e^{-\frac{S_2}{2} Y^2} \dots e^{-\frac{S_m}{m} Y^m} \dots$$

La identificación, teniendo en cuenta que $e^U = \sum_{k \geq 0} \frac{U^k}{k!}$, conduce a

$$(12) \quad p_m = (-1)^m \sigma_m = \sum_{r_1 + 2r_2 + \dots + mr_m = m} \frac{(-1)^{r_1 + r_2 + \dots + r_m}}{r_1! r_2! \dots r_m! 2^{r_2} 3^{r_3} \dots m^{r_m}} S_1^{r_1} S_2^{r_2} \dots S_m^{r_m}.$$

A modo de ejemplo, utilizamos (11) para calcular S_5 cuando $n \geq 5$. Si

$$r_1 + 2r_2 + \dots + nr_n = 5,$$

se tiene $r_n = 0$ para $n \geq 6$. Las soluciones enteras, positivas o nulas de la ecuación $r_1 + 2r_2 + 3r_3 + 4r_4 + 5r_5 = 5$, los valores de $r = \sum r_i$ y del coeficiente $\frac{(r_1 + r_2 + r_3 + r_4 + r_5 - 1)!}{r_1! r_2! r_3! r_4! r_5!}$, se hallan consignados en la tabla siguiente:

r_1	r_2	r_3	r_4	r_5	r	$\frac{(r-1)!}{\prod r_i!}$
0	0	0	0	1	1	1
0	1	1	0	0	2	1
1	0	0	1	0	2	1
1	2	0	0	0	3	1
2	0	1	0	0	3	1
3	1	0	0	0	4	1
5	0	0	0	0	5	$\frac{1}{5}$

de donde (volviendo a los σ_i):

$$\frac{S_5}{5} = \frac{1}{5} \sigma_1^5 - \sigma_1^3 \sigma_2 + \sigma_1^2 \sigma_3 + \sigma_1 \sigma_2^2 - \sigma_1 \sigma_4 - \sigma_2 \sigma_3 + \sigma_5.$$

§ V.3 ECUACIONES DE SEGUNDO Y DE TERCER GRADO

● Aquí el cuerpo base es \mathbf{C} .

Consideremos la ecuación algebraica $P(x) = 0$ (P polinomio de grado n). Sabemos (T. de d'Alembert) que admite n raíces (distintas o no). Escribimos

$$P(x) = x^n + a_1 x^{n-1} + \dots + a_n.$$

El problema más natural consiste en buscar fórmulas «explícitas» que nos den los valores de las raíces en función de a_1, a_2, \dots, a_n . (En otros términos, intentar expresar las raíces en función de los parámetros (a_i) .) Desde la Antigüedad se sabe resolver la ecuación

$$x^2 + ax + b = 0$$

poniendo $\Delta = \frac{a^2 - 4b}{4}$ se obtiene la ecuación $\left(x + \frac{a}{2}\right)^2 = \Delta$, y el problema equivale a extraer las raíces cuadradas de Δ .

Recordemos que siempre es posible obtener fórmulas *en que sólo intervengan radicales reales*. En efecto, si hacemos $\Delta = \alpha + i\beta$, α y β reales (i , número complejo de cuadrado -1) y $z = x + iy$, x e y reales. La ecuación $z^2 = \Delta$ equivale al sistema

$$\begin{aligned} (1) \quad & x^2 - y^2 = \alpha, \\ (2) \quad & x^2 + y^2 = \sqrt{\alpha^2 + \beta^2}, \\ (3) \quad & 2xy = \beta. \end{aligned}$$

(1) y (2) determinan $x^2 = \frac{1}{2}(\sqrt{\alpha^2 + \beta^2} + \alpha)$, $y^2 = \frac{1}{2}(\sqrt{\alpha^2 + \beta^2} - \alpha)$, y a priori, resultan 4 valores posibles de z . Pero (3) liga los signos de x e y , de ahí que sólo 2 sean soluciones.

La ecuación de tercer grado

$$(4) \quad x^3 + ax^2 + bx + c = 0$$

también se puede resolver «por radicales». La resolución no se abordó hasta el siglo XIII, y no se esclareció completamente hasta Cardan (1501-1576).

Estudio de (4)

Reemplazando x por $x - \frac{a}{3}$ se transforma en

$$(5) \quad x^3 + px + q = 0.$$

La observación fundamental (Lagrange) es ésta: hacemos $j = e^{\frac{2\pi i}{3}} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.

Entonces la función

$$\varphi = (x_1 + jx_2 + j^2 x_3)^3$$

sólo toma 2 valores si se efectúan todas las permutaciones de x_1, x_2, x_3 , que son φ_1 y φ_2 . Observemos que φ_1 y φ_2 son funciones *alternadas* de x_1, x_2, x_3 , es decir, invariantes por las tres permutaciones circulares de x_1, x_2, x_3 .

De esto resulta que los coeficientes de la ecuación

$$(6) \quad (X - \varphi_1)(X - \varphi_2) = 0,$$

a saber, $\varphi_1 + \varphi_2$, y $\varphi_1 \varphi_2$, son *simétricos* en x_1, x_2, x_3 . Luego (§ 1) $\varphi_1 + \varphi_2$ y $\varphi_1 \varphi_2$ se pueden expresar con la ayuda de

$$\sigma_1 = x_1 + x_2 + x_3, \quad \sigma_2 = x_2 x_3 + x_3 x_1 + x_1 x_2, \quad \sigma_3 = x_1 x_2 x_3.$$

Si hacemos que x_1, x_2, x_3 sean las raíces de (5), entonces $\sigma_1 = 0, \sigma_2 = p, \sigma_3 = -q$, y obtenemos los coeficientes de (6) en función de p y q ; (5) es de grado 2, y su resolución dará φ_1 y φ_2 .

Luego, conociendo φ_1 y φ_2 , el cálculo de x_1, x_2, x_3 equivale a la extracción de raíces cúbicas: sean u, v raíces cúbicas de φ_1, φ_2 tales que $uv = -3p$; los x_i vienen dados entonces por el sistema

$$(R) \quad \begin{cases} x_1 + jx_2 + j^2x_3 = u \\ x_1 + j^2x_2 + jx_3 = v \\ x_1 + x_2 + x_3 = 0 \end{cases}$$

$$\text{de donde} \quad x_1 = \frac{1}{3}(u + v), \quad x_2 = \frac{1}{3}(ju + j^2v), \quad x_3 = \frac{1}{3}(j^2u + jv)$$

(la elección de otras determinaciones para las raíces cúbicas u y v conduciría a permutaciones sobre x_1, x_2, x_3).

Vemos, pues, que la resolución de (5) se reduce a la de (6). Por esta razón a (6) se le designa como *una resolvente de (5)*.

Para calcular efectivamente φ_1 y φ_2 , no construiremos (6) (ello conduciría a cálculos pesados).

En efecto, φ_1 y φ_2 son alternadas, luego $\varphi_1 + \varphi_2$ es simétrica y $\varphi_1 - \varphi_2 = \Phi$ es alternada, y además Φ se transforma en $-\Phi$ por medio de las permutaciones impares. Pongamos $\delta = (x_3 - x_1)(x_3 - x_2)(x_2 - x_1)$. Sabemos (Cap. IV final del § 7) que existe un polinomio simétrico S_2 tal que

$$\varphi_1 - \varphi_2 = 2\delta S_2.$$

Poniendo $\varphi_1 + \varphi_2 = 2S_1$, se obtiene:

$$(7) \quad \begin{cases} \varphi_1 = S_1 + \delta S_2 \\ \varphi_2 = S_1 - \delta S_2, \end{cases}$$

en donde S_1, S_2 son polinomios simétricos determinados de forma única por (7) (por razones del grado, S_2 es una constante).

Ahora bien, el desarrollo de $\varphi_1 = (x_1 + jx_2 + j^2x_3)^3$ da:

$$\begin{aligned} \varphi_1 &= x_1^3 + x_2^3 + x_3^3 + 3j(x_1^2x_2 + x_2^2x_3 + x_3^2x_1) + 3j^2(x_1x_2^2 + x_2x_3^2 + x_3x_1^2) + \\ &\quad + 6x_1x_2x_3 \\ (8) \quad &= s_3 - \frac{3}{2}(\sum x_1^2x_2) + 3i\frac{\sqrt{3}}{2} \times \\ &\quad \times (x_1^2x_2 + x_2^2x_3 + x_3^2x_1 - x_1x_2^2 - x_2x_3^2 - x_3x_1^2) + 6\sigma_3, \end{aligned}$$

(haciendo $s_k = x_1^k + x_2^k + x_3^k$). Un cálculo fácil (cf. § 1) da

$$\begin{cases} \sum x_1^2x_2 = \sigma_1\sigma_3 - 3\sigma_3 = 3q \\ s_3 = 3\sigma_3 = -3q \\ \sigma_3 = -q. \end{cases}$$

En (8) el coeficiente de $3i\frac{\sqrt{3}}{2}$ es alternado, ya que los otros términos son simétricos. Puesto que es de grado 3, es $\pm \delta$, y se comprueba que es $-\delta$.

Se concluye

$$(9) \quad \begin{cases} \varphi_1 = -\frac{27}{2}q - 3i\frac{\sqrt{3}}{2}\delta, \\ \varphi_2 = -\frac{27}{2}q + 3i\frac{\sqrt{3}}{2}\delta. \end{cases}$$

Nos vemos obligados a calcular δ , lo que es relativamente simple. Se tiene

$$\delta^2 = \prod_{i < j} (x_i - x_j)^2 \quad (\text{discriminante de (5); cf. § VI.4})$$

δ^2 es simétrico de peso 6 y de orden 4. Como $\sigma_1 = \sum x_i = 0$, se tiene

$$\delta^2 = A\sigma_2^3 + B\sigma_3^2 = Ap^3 + Bq^2,$$

en donde A y B son constantes numéricas (cf. § 1).

Calculando explícitamente δ^2 para las ecuaciones $x(x^2 - 1) = 0$ y $x^3 - 1 = 0$ se obtiene

$$A = -4 \text{ y } B = -27 \text{ de donde } \delta^2 = -4p^3 - 27q^2.$$

Convenimos en designar por $\sqrt[m]{a}$ a una de las raíces n -ésimas del número complejo a .

Volviendo al sistema (R), las tres raíces se obtienen entonces en la forma

$$(10) \quad \begin{cases} x_1 = \frac{1}{3}(u+v) \\ x_2 = \frac{1}{2}(ju + j^2v) \\ x_3 = \frac{1}{3}(j^2u + jv) \end{cases} \quad \text{con} \quad \begin{cases} u = \sqrt[3]{-\frac{27q}{2} + 3i\frac{\sqrt{3}}{2}\sqrt{-4p^3 - 27q^2}} \\ v = \sqrt[3]{-\frac{27q}{2} - 3i\frac{\sqrt{3}}{2}\sqrt{-4p^3 - 27q^2}} \end{cases}$$

(fórmulas de Cardan).

En la práctica, no se rehace esta teoría, sino que se buscan desde el principio las raíces de (5) que son de la forma $x_1 = \alpha + \beta$, $x_2 = \alpha j + \beta j^2$, $x_3 = \alpha j^2 + \beta j$. Se tiene la identidad

$$[X - (\alpha + \beta)][X - (\alpha j + \beta j^2)][X - (\alpha j^2 + \beta j)] = X^3 - 3\alpha\beta X - (\alpha^3 + \beta^3).$$

La condición para que esta ecuación sea equivalente a (5) es entonces

$$\begin{aligned} p &= -3\alpha\beta \\ -q &= \alpha^3 + \beta^3. \end{aligned}$$

Luego, p es una de las tres raíces cúbicas de $-27\alpha^3\beta^3$. Haciendo $U = \alpha^3$ y $V = \beta^3$, U y V se hallan definidas por $U + V = -q$, $UV = -p^3/27$, y son raíces de la ecuación resolvente

$$X^2 + qX - \frac{p^3}{27} = 0,$$

que es precisamente la transformada de (6) por $X = 27Y$, ya que por los cálculos realizados antes es $\varphi_1 + \varphi_2 = -27q$ y $\varphi_1\varphi_2 = -27p^3$.

Estudio de (4) cuando p y q son reales

Consideremos de nuevo las fórmulas anteriores (10), que se pueden escribir:

$$u = \sqrt[3]{-\frac{27}{2}q + 3\frac{\sqrt{3}}{2}\sqrt{\Delta}}, \quad v = \sqrt[3]{-\frac{27}{2}q - 3\frac{\sqrt{3}}{2}\sqrt{\Delta}},$$

con $\Delta = 4p^3 + 27q^2$; y supongamos $\Delta \neq 0$.

Si $\Delta > 0$, elegido $\sqrt{\Delta}$ (por ejemplo, > 0), se obtiene un par (u, v) de números reales, único, y $\Delta \neq 0 \Rightarrow u \neq v$. Luego los números

$$x_2 = \frac{1}{3} (ju + \bar{j}v) \quad \text{y} \quad x_3 = \frac{1}{3} (\bar{j}u + jv)$$

son *no reales* y conjugados. En este caso, sólo una de las raíces es real, y esta raíz se expresa con la ayuda de *radicales reales*:

$$(11) \quad x_1 = \frac{1}{3} (u + v).$$

Si x_1 ha sido determinada, (11) da relaciones no triviales entre irracionales.

En la Edad Media estas relaciones constituían el punto de partida de los estudios acerca de la ecuación (5). Por ejemplo, sea $\gamma \in \mathbf{Q}$, tal que el número $\Delta = \frac{4(\gamma-1)^3}{27} + \gamma^2$ (relativo a la ecuación $(x-1)(x^2+x+\gamma)=0$) sea > 0 . Escribiendo (11) se obtendrá una de estas relaciones.

Así, para $\gamma = 2$, (11) conduce a

$$\sqrt{3} = \sqrt[3]{2\sqrt{7} + 3\sqrt{3}} - \sqrt[3]{2\sqrt{7} - 3\sqrt{3}}$$

(relación que se puede intentar verificar directamente).

Si $\Delta < 0$, los números u^3 y v^3 son *no reales y conjugados*. Siempre es posible elegir u y v convenientemente para que sean conjugados (necesariamente no reales).

Las fórmulas (9) se convierten entonces en

$$x_1 = \frac{1}{3} (u + \bar{u}), \quad x_2 = \frac{1}{3} (ju + \bar{j}u), \quad x_3 = \frac{1}{3} (j^2u + \bar{j}^2u),$$

lo que demuestra que *las 3 raíces son reales*. Pero este caso es precisamente aquel en que $\sqrt{\Delta}$ no puede representar un número real. En una época en la que se desconocía el uso de los números complejos, se había observado que, en el caso en que $\sqrt{\Delta}$ «carecía de sentido», (5) poseía 3 raíces reales; fue esta observación la que determinó a Cardan a calcular utilizando números «ideales», que se convertirían en los números complejos.

Hay más en este caso $\Delta < 0$. Las fórmulas que dan u y v sólo contienen radicales. Pero bajo el radical cúbico se halla el número no real $-\frac{27}{2}q + 3i\frac{\sqrt{3}}{2}\sqrt{-\Delta}$. Si se intenta expresar $\operatorname{Re}(u)$ e $\operatorname{Im}(u)$ con la ayuda exclusiva de radicales reales, nos vemos obligados a resolver ecuaciones de grado 3, *de las que se sabe de antemano que todas las raíces son reales*; y se recae indefinidamente en el mismo pro-

blema. De hecho, con la ayuda de la teoría de Galois, se puede demostrar que *es imposible en este caso expresar por medio de radicales reales las raíces de (5)* ⁽¹⁾. (Pero este resultado sobrepasa el nivel de nuestra obra.)

Trisección del ángulo

El viejo problema de dividir un ángulo en 3 ángulos iguales se puede expresar de la forma siguiente: conociendo $a = \cos 3\varphi$, hallar $u = \cos \varphi$.

Con ayuda de la siguiente relación $\cos 3\varphi = 4\cos^3 \varphi - 3\cos \varphi$, vemos que el problema consiste en resolver la ecuación de tercer grado

$$(12) \quad 4u^3 - 3u - a = 0.$$

Un estudio de Análisis prueba que todas las raíces de (12) son reales. Por otra parte, $a \in]-1, +1[$ y $\Delta = \frac{27}{16}(a^2 - 1) < 0$, por lo que, en general, no es posible resolver (12) mediante radicales reales.

Inversamente, esta interpretación permite utilizar las tablas de las funciones circulares para resolver numéricamente ciertas ecuaciones de grado 3.

De hecho, si

$$(13) \quad x^3 + px + q = 0, \text{ con } p, q \text{ reales}$$

posee 3 raíces reales, se tiene $p < 0$. Haciendo $\lambda = \sqrt{-\frac{4p}{3}}$ y $x = \lambda y$, (13) se transforma en

$$4y^3 - 3y = \frac{3q}{p\lambda},$$

y, haciendo $y = \cos \varphi$, queda

$$(14) \quad \cos 3\varphi = \frac{3q}{p\lambda}.$$

La relación $4p^3 + 27q^2 < 0$ implica $\left| \frac{3q}{p\lambda} \right| < 1$, y entonces (14) puede dar 3φ recurriendo a las tablas de los cosenos, y por ende φ e $y = \cos \varphi$.

⁽¹⁾ Por esta razón, este caso de la ecuación de tercer grado fue llamado «casus irreducibilis».

§ V.4 ECUACIÓN DE CUARTO GRADO

Una circunstancia análoga a la del principio de § 2 permite resolver completamente las ecuaciones de grado 4. La ecuación

$$(1) \quad x^4 + ax^3 + bx^2 + cx + d = 0$$

admite 4 raíces $(x_i)_{i=1,2,3,4}$ pero existen funciones racionales de las x_i que toman 3 valores tras una permutación cualquiera de las x_i .

Es, por ejemplo, el caso de la función

$$y = x_1 x_2 + x_3 x_4$$

que, después de haber permutado las x_i , toma los tres valores

$$y_1 = y, \quad y_2 = x_1 x_3 + x_2 x_4 \quad \text{y} \quad y_3 = x_1 x_4 + x_2 x_3.$$

Se deduce que los coeficientes de la ecuación

$$(R) \quad (U - y_1)(U - y_2)(U - y_3) = 0$$

son simétricos en x_1, x_2, x_3, x_4 .

Se pueden calcular en función de a, b, c, d , racionalmente. Como (R) es de grado 3 su resolución es posible y da las y_i .

El cálculo completo de las x_i se hace entonces fácilmente resolviendo el sistema

$$(2) \quad \begin{cases} x_1 x_2 + x_3 x_4 & = y_1 & x_1 x_2 x_3 x_4 = d \\ x_1 x_3 + x_2 x_4 & = y_2 \\ x_1 x_4 + x_2 x_3 & = y_3 \\ x_1 + x_2 + x_3 + x_4 & = -a \end{cases}$$

(Por ejemplo, sumando las dos primeras relaciones (2) se obtiene

$$(x_1 + x_4)(x_2 + x_3) = y_1 + y_2$$

que con la última relación (2), determina

$$A_1 = x_1 + x_4 \quad \text{y} \quad A_2 = x_2 + x_3;$$

luego las relaciones

$$x_1 x_4 + x_2 x_3 = y_3 \quad \text{y} \quad x_1 x_2 x_3 x_4 = d$$

determinan $B_1 = x_1 x_4$ y $B_2 = x_2 x_3$. Conociendo A_1, B_1 se deduce x_1 y x_4 , y con A_2, B_2 se obtienen análogamente x_2 y x_3 .)

Vamos a calcular (R) para la ecuación

$$(3) \quad x^4 + ax^2 + bx + c = 0, \quad c \neq 0 \quad (\text{pues } c = 0 \text{ conduce al 3.º grado}).$$

Observemos que (1) se transforma en (3) si reemplazamos x por $x - \frac{a}{4}$, luego resolver (3) equivale a tratar el caso general.

Las funciones simétricas σ_i de las raíces de (3) están dadas por

$$\sigma_1 = 0, \quad \sigma_2 = a, \quad \sigma_3 = -b, \quad \sigma_4 = c$$

Se obtiene sucesivamente:

$$y_1 + y_2 + y_3 = \sigma_2,$$

$$y_2 y_3 + y_3 y_1 + y_1 y_2 = \sum x_1^2 x_2 x_3,$$

$$y_1 y_2 y_3 = \sigma_4 (\sum x_i^2) + \sum x_1^2 x_2^2 x_3^2.$$

Según § 1, se sabe calcular $\sum x_1^2 x_2 x_3 = \sigma_1 \sigma_3 - 4\sigma_4$.

Finalmente $\sum x_1^2 x_2^2 x_3^2 = \sigma_4^2 \sum \frac{1}{x_i^2}$. La ecuación cuyas raíces son las $\frac{1}{x_i}$ es evidentemente:

$$cx^4 + bx^3 + ax^2 + 1 = 0,$$

de donde, si designamos por σ'_i las funciones simétricas de las raíces de esta ecuación,

$$\sum \frac{1}{x_i^2} = \sigma_i'^2 - 2\sigma_2', \quad \text{con} \quad \sigma_1' = -\frac{b}{c}, \quad \sigma_2' = \frac{a}{c}.$$

Recapitulando, se obtiene

$$y_1 + y_2 + y_3 = a, \quad y_2 y_3 + y_3 y_1 + y_1 y_2 = -4c, \quad y_1 y_2 y_3 = b^2 - 4ac.$$

De donde:

V.4.1 Para la ecuación (3): $x^4 + ax^2 + bx + c = 0$, la resolvente (R) correspondiente a la función $y = x_1 x_2 + x_3 x_4$ es

$$\parallel \parallel \quad y^3 - ay^2 - 4cy + 4ac - b^2 = 0.$$

He aquí como Ferrari (1522-1565) resolvió (3):

Se escribe (3) en la forma

$$x^4 = -ax^2 - bx - c, \text{ o sea (para todo } y \in \mathbf{C}):$$

$$x^4 + x^2 y + \frac{1}{4}y^2 = (y - a)x^2 - bx + \frac{1}{4}y^2 - c$$

$$(4) \quad \left(x^2 + \frac{1}{2}y\right)^2 = (y - a)x^2 - bx + \frac{1}{4}y^2 - c.$$

Se intenta obtener, en ambos miembros, cuadrados de trinomios y para ello se busca y de tal forma que el segundo miembro de (4) sea un cuadrado. Para que el segundo miembro de (4) sea de la forma $(mx + n)^2$ es necesario y suficiente que el discriminante $b^2 - 4\left(\frac{1}{4}y^2 - c\right)(y - a)$ sea nulo. Desarrollando esta relación nos da

$$y^3 - ay^2 - 4cy + 4ac - b^2 = 0,$$

es decir, precisamente (R).

Explicación. Si y_1 es una raíz de (R), (4) implica

$$\left. \begin{array}{l} x^2 + \frac{1}{2}y_1 = mx + n \\ x^2 + \frac{1}{2}y_1 = -mx - n \end{array} \right\} \Rightarrow \left. \begin{array}{l} x_1 x_2 = \frac{1}{2}y_1 - n \\ x_3 x_4 = \frac{1}{2}y_1 + n \end{array} \right\} \Rightarrow y_1 = x_1 x_2 + x_3 x_4.$$

luego los valores de y que hacen del segundo miembro de (4) un cuadrado perfecto son exactamente los números y_1, y_2, y_3 definidos antes.

Este método de Ferrari es, al mismo tiempo, un método práctico muy elegante para resolver (3).

Nota. Se tiene $y_1 - y_2 = (x_1 - x_4)(x_2 - x_3)$, etc.

Luego $\prod_{i < j} (x_i - x_j)^2 = \prod_{i < j} (y_i - y_j)^2$: el discriminante de (3) es igual al de su resolvente (lo que proporciona un método para calcularlo).

Otro método de resolución de (3)

Se buscan α, β, γ tales que

$$x^4 + ax^2 + bx + c = (x^2 + \alpha x + \beta)(x^2 - \alpha x + \gamma),$$

de donde las condiciones $c = \beta\gamma$, $a = \beta + \gamma - \alpha^2$, $b = \alpha(\gamma - \beta)$. Si $b \neq 0$, $\alpha \neq 0$, de

$$\beta + \gamma = a + \alpha^2 \quad \text{y} \quad \beta - \gamma = -\frac{b}{\alpha}$$

se obtiene

$$\beta = \frac{1}{2} \left(a + \alpha^2 - \frac{b}{\alpha} \right) \quad \text{y} \quad \gamma = \frac{1}{2} \left(a + \alpha^2 + \frac{b}{\alpha} \right).$$

La condición $c = \beta\gamma$ se escribe pues

$$(5) \quad c = \frac{1}{4} \left[(a + \alpha^2)^2 - \frac{b^2}{\alpha^2} \right].$$

Haciendo $Z = \alpha^2$ se obtiene la resolvente

$$(R') \quad Z^3 + 2aZ^2 + (a^2 - 4c)Z - b^2 = 0,$$

que es diferente de (R).

Sea z_1 una raíz de (R'); (3) se resuelve con la ayuda de dos ecuaciones trinomias:

$$x^2 + \sqrt{z_1}x + \frac{1}{2} \left(a + z_1 - \frac{b}{\sqrt{z_1}} \right) = 0, \quad \text{de donde} \quad x_1 + x_2 = -\sqrt{z_1}$$

$$x^2 - \sqrt{z_1}x + \frac{1}{2} \left(a + z_1 + \frac{b}{\sqrt{z_1}} \right) = 0, \quad \text{de donde} \quad x_3 + x_4 = \sqrt{z_1}.$$

de donde se deduce

$$z_1 = \frac{1}{4} (x_1 + x_2 - x_3 - x_4)^2.$$

Luego, por permutación de las x_i , la función $z = \frac{1}{4} (x_1 + x_2 - x_3 - x_4)^2$ toma sólo 3 valores. Por lo tanto, nos habríamos podido servir de z exactamente igual que de y , y la resolvente que corresponde a z es (R'). La verificación directa de este hecho es pesada.

Ecuaciones particulares

Quede claro que estas observaciones acerca de la ecuación de 4.º grado se refieren a la *ecuación general*. En ciertos casos particulares, la resolución es inmediata, así, recordemos que la ecuación bicuadrada:

$$X^4 + aX^2 + b = 0$$

se resuelve haciendo $X^2 = U$. Vamos a calcular, a modo de ejemplo, las funciones circulares del ángulo $\frac{\pi}{8}$.

Las raíces de la ecuación $X^8 + 1 = 0$ son $e^{\pm i\frac{\pi}{8}}, e^{\pm 3i\frac{\pi}{8}}, e^{\pm 5i\frac{\pi}{8}}, e^{\pm 7i\frac{\pi}{8}}$; descomponemos $X^8 + 1$ en $\mathbf{R}[X]$:

$$X^8 + 1 = (X^4 + 1)^2 - 2X^4 = (X^4 - \sqrt{2}X^2 + 1)(X^4 + \sqrt{2}X^2 + 1)$$

las raíces de cada una de las ecuaciones bicuadradas que figura en el segundo miembro son opuestas dos a dos. Como la parte real de $(e^{i\frac{\pi}{8}})^2 = e^{i\frac{\pi}{4}}$ es $\frac{\sqrt{2}}{2}$, vemos que los números $e^{\pm i\frac{\pi}{8}}, e^{\pm 7i\frac{\pi}{8}}$ son las raíces de la ecuación bicuadrada:

$$X^4 - \sqrt{2}X^2 + 1 = 0.$$

Descomponemos el primer miembro en $\mathbf{R}[X]$:

$$\begin{aligned} X^4 - \sqrt{2}X^2 + 1 &= (X^2 + 1)^2 - (2 + \sqrt{2})X^2 = \\ &= (X^2 - \sqrt{2 + \sqrt{2}}X + 1)(X^2 + \sqrt{2 + \sqrt{2}}X + 1). \end{aligned}$$

La parte real de $e^{i\frac{\pi}{8}}$ es ≥ 0 , luego $e^{\pm i\frac{\pi}{8}}$ son las raíces de

$$X^2 - \sqrt{2 + \sqrt{2}}X + 1 = 0.$$

Se deduce:

$$2 \cos \frac{\pi}{8} = \sqrt{2 + \sqrt{2}}, \quad 2 \sin \frac{\pi}{8} = \sqrt{2 - \sqrt{2}}.$$

En los ejercicios se encontrará el cálculo completo de los

$$\cos \frac{k\pi}{2^{n-1}}, \quad 0 \leq k \leq 2^{n-1}.$$

§ V.5 ECUACIONES DE GRADO ≥ 5

A priori podríamos preguntarnos por qué los métodos expuestos en el § 2 no son aplicables a la ecuación más general. El problema consiste en hallar funciones racionales de x_1, \dots, x_n que tomen más de 2 y menos de $n - 1$ valores tras cualquier permutación de las x_i . Desgraciadamente, para $n \geq 5$ los coeficientes de tales funciones vienen determinados por ecuaciones de grado $\geq n$; ello se debe al hecho de que para $n \geq 5$ el grupo simétrico \mathfrak{S}_n posee bruscamente una estructura más

pobre, «irreducible» en cierto sentido (con precisión, el grupo alternado \mathcal{A}_n no admite subgrupos normales propios para $n \geq 5$). A partir de estas observaciones, Galois y Abel han demostrado que para $n \geq 5$ no se puede, en general, dar fórmulas de resolución por medio de radicales. ⁽¹⁾

He aquí ciertos casos en los cuales es posible reducir el grado de una ecuación, hasta resolverla.

Ecuaciones recíprocas

DEFINICIÓN V.5.1

~ A la ecuación

$$(1) \quad a_0 x^n + a_1 x^{n-1} + \cdots + a_n = 0$$

~ se le llama **recíproca de primera especie** si

$$a_k = a_{n-k} \quad \left(k \leq \frac{n}{2} \right);$$

~ **recíproca de segunda especie** si $a_k + a_{n-k} = 0 \quad \left(k \leq \frac{n}{2} \right).$

Supongamos que (1) es recíproca de segunda especie y n arbitrario: $x = 1$ es una raíz, y el primer miembro de (1) se puede escribir $(x - 1)g(x)$, en donde la ecuación $g(x) = 0$ es recíproca de primera especie.

Supongamos que (1) es recíproca de primera especie y n impar: $x = -1$ es una raíz, y el primer miembro de (1) se escribe $(x + 1)g(x)$, en donde la ecuación $g(x) = 0$ es recíproca de primera especie, y de grado par. Podemos enunciar, pues:

V.5.1 Cuando una ecuación recíproca no admite ni a 1 ni a -1 por raíces, es
|| de primera especie y de grado par.

En la práctica, para resolver (1) se empieza por quitarle las raíces ± 1 . Estudiemos la ecuación recíproca de primera especie y de grado par $2n$, o sea:

$$\begin{aligned} (2) \quad a_0 x^{2n} + a_1 x^{2n-1} + \cdots + a_{n-1} x^{n+1} + a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 &= 0, \\ a_0 &\neq 0. \end{aligned}$$

⁽¹⁾ (Nota del Autor). En el caso de una ecuación de grado 5 que puede resolverse por medio de radicales, el autor ha dado su método explícito de cálculo de las raíces, que conduce a fórmulas análogas a las «de Cardan». Véase *Bulletin des Sciences Mathématiques*, n.º 100, octubre 1976, *Sur la Résolution explicite des équations de degré 5 quand elles sont résolubles par radicaux*.

$x = 0$ no es raíz, por lo tanto (2) es equivalente a

$$(3) \quad a_0 \left(x^n + \frac{1}{x^n} \right) + a_1 \left(x^{n-1} + \frac{1}{x^{n-1}} \right) + \cdots + a_{n-1} \left(x + \frac{1}{x} \right) + a_n = 0.$$

Hagamos

$$U = x + \frac{1}{x} \text{ y } S_n = x^n + \frac{1}{x^n} \text{ para } n \geq 0 \quad (S_0 = 2, S_1 = U), \quad US_n = S_{n+1} + S_{n-1}.$$

Los S_n se determinan entonces mediante la relación de recurrencia

$$(4) \quad \begin{cases} S_n = US_{n-1} - S_{n-2} & (n \geq 2) \\ S_0 = 2, \quad S_1 = U \end{cases}$$

que permite calcular los S_n , paso a paso, como polinomios en U . Por recurrencia, se ve que S_n es un polinomio de grado n , cuyo término de mayor grado es U^n . En otras palabras, (2) admite una resolvente de grado n en U . A cada raíz U_0 de esta resolvente corresponden dos raíces x_1 y x_2 de (2), inversas una de otra y dadas por la ecuación:

$$U_0 = x + \frac{1}{x}, \text{ o sea } x^2 - U_0 x + 1 = 0.$$

Enunciaremos:

V.5.2 Una ecuación recíproca de primera especie y de grado n par se transforma, || por medio de la transformación $U = x + \frac{1}{x}$, en una ecuación de grado $n/2$.

Es fácil dar la expresión general de S_n . Más adelante veremos (§ XI.4) que toda relación lineal de recurrencia con tres términos

$$\begin{cases} u_n = \alpha u_{n-1} + \beta u_{n-2} \\ u_1 = a_1, \quad u_0 = a_0 \end{cases}$$

admite como solución una sucesión (u_n) definida por

$$u_n = a_1 P_n + a_0 Q_n \quad (n \geq 2)$$

$$Q_{n+1} = \beta P_n \quad (n \geq 2) \quad \text{y} \quad P_n = \alpha^{n-1} + \sum_{1 \leq k \leq \frac{n-1}{2}} \binom{n-1-k}{k} \alpha^{n-1-2k} \beta^k.$$

Apliquemos esta relación a (4): $\alpha = U$, $\beta = -1$, $a_1 = U$, $a_0 = 2$. Se obtiene:

$$S_n = UP_n + 2Q_n$$

$$(5) \quad \begin{aligned} P_n &= U^{n-1} + \sum_{1 \leq k \leq \frac{n-1}{2}} \binom{n-1-k}{k} (-1)^k U^{n-1-2k}, \\ Q_n &= -P_{n-1} = -U^{n-2} - \sum_{1 \leq l \leq \frac{n-2}{2}} \binom{n-2-l}{l} (-1)^l U^{n-2-2l}. \end{aligned}$$

S_n posee evidentemente la paridad de n .

Aplicación. En lo que precede, si reemplazamos x por $e^{i\theta}$ es:

$$U = 2 \cos \theta, \quad S_n = 2 \cos n\theta.$$

Las fórmulas (5) nos dan directamente la expresión de $\cos n\theta$ como polinomio respecto de $\cos \theta$, expresión encontrada ya en la página 160.

Ejemplos de ecuaciones recíprocas

1) Raíces quintas de 1:

Las raíces quintas de 1 vienen dadas por $x^5 - 1 = 0$, lo que equivale a

$$(6) \quad (x - 1)(x^4 + x^3 + x^2 + x + 1) = 0.$$

La ecuación (6): $x^4 + x^3 + x^2 + x + 1 = 0$, admite las raíces $e^{\pm \frac{2\pi i}{5}}$, $e^{\pm \frac{4\pi i}{5}}$. Haciendo $S_2 = x^2 + \frac{1}{x^2}$ y $U = x + \frac{1}{x}$, (6) se escribe

$$S_2 + U + 1 = 0,$$

con lo que, teniendo en cuenta que $S_2 = U^2 - 2$, la resolvente en U es:

$$(7) \quad U^2 + U - 1 = 0.$$

Las raíces de (7) son

$$\begin{aligned} x_1 + \frac{1}{x_1} &= e^{\frac{2\pi i}{5}} + e^{-\frac{2\pi i}{5}} = 2 \cos \frac{2\pi}{5}, \\ x_2 + \frac{1}{x_2} &= e^{\frac{4\pi i}{5}} + e^{-\frac{4\pi i}{5}} = 2 \cos \frac{4\pi}{5}. \end{aligned}$$

Puesto que $\cos \frac{2\pi}{5} > 0$ y $\cos \frac{4\pi}{5} < 0$, (7) proporciona valores con radicales de estos números:

$$\cos \frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}, \quad \cos \frac{4\pi}{5} = -\frac{\sqrt{5}+1}{4}.$$

Con la ayuda de la relación $\cos^2 \theta + \sin^2 \theta = 1$, se deduce:

$$\sin \frac{2\pi}{5} = \frac{\sqrt{10+2\sqrt{5}}}{4}, \quad \sin \frac{4\pi}{5} = \frac{\sqrt{10-2\sqrt{5}}}{4},$$

$$e^{\frac{2\pi i}{5}} = \frac{\sqrt{5}-1}{4} + i \frac{\sqrt{10+2\sqrt{5}}}{4}, \quad e^{\frac{4\pi i}{5}} = -\frac{\sqrt{5}+1}{4} + i \frac{\sqrt{10-2\sqrt{5}}}{4}.$$

Como ejercicio, el lector podrá deducir de estos resultados todos los elementos notables de un pentágono y de un decágono regular, así como una construcción por medio de regla y compás del lado de estos polígonos, conociendo su círculo circunscrito.

2) Raíces séptimas de 1.

Dadas por $x^7 - 1 = 0$, o sea por $(x-1)(x^6 + x^5 + x^4 + x^3 + x^2 + x + 1) = 0$. Las raíces de

$$(8) \quad x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 = 0$$

son $e^{\pm \frac{2\pi i}{7}}$, $e^{\pm \frac{4\pi i}{7}}$, $e^{\pm \frac{6\pi i}{7}}$. Para obtener la resolvente en $U = x + \frac{1}{x}$, se divide por x^3 ,

y se obtiene $S_2 = x^2 + \frac{1}{x^2}$ y $S_3 = x^3 + \frac{1}{x^3}$ por medio de los cálculos siguientes:

$$S_2 = U^2 - 2, \quad S_3 = US_2 - S_1 = U^3 - 3U,$$

con lo que se obtiene la resolvente

$$(9) \quad U^3 + U^2 - 2U - 1 = 0.$$

Las 3 raíces de (9) son $2 \cos \frac{2\pi}{7}$, $2 \cos \frac{4\pi}{7}$ y $2 \cos \frac{6\pi}{7}$; son reales (es fácil comprobarlo directamente). Luego (§ 2) no es posible expresarlas por radicales reales, a

menos que una de ellas no sea racional. Pero si (9) posee la raíz racional $U = \frac{a}{b}$, con a y b enteros primos entre sí, se tiene:

$$a^3 + a^2 b - 2 ab^2 - b^3 = 0,$$

de donde se deduce que a divide a b^3 y b divide a a^3 , luego $a = \pm b = \pm 1$. Pero (9) no admite ni a 1 ni a -1 por raíz, luego (9) carece de raíces racionales, y *no es posible expresar el lado del heptágono regular (inscrito en el círculo unidad) con la ayuda de radicales reales.*

Capítulo VI

Eliminación

La teoría de la eliminación constituye la base de la teoría de los sistemas de ecuaciones algebraicas. El problema general es el siguiente: dado un sistema de ecuaciones algebraicas $f_i = 0$ ($1 \leq i \leq p$), en donde f_i es un polinomio con n variables, y con coeficientes en un cuerpo K , hallar las condiciones necesarias y suficientes, que relacionen los *coeficientes* de los f_i , para que el sistema admita, por lo menos, una solución.

Observemos que este problema se resolverá por métodos directos en el capítulo X, § 4 en el caso en que todos los f_i sean de grado 1. Fue estudiando este caso particular que Leibnitz descubrió los determinantes.

Daremos una respuesta satisfactoria al problema de la eliminación en el caso de un sistema de dos ecuaciones de grado arbitrario.

Para ello utilizaremos ciertos resultados de la teoría de los determinantes y de las ecuaciones lineales desarrollados en el capítulo X. Es, pues, preferible no abordar la lectura de este capítulo (relativamente más difícil) en tanto no se haya realizado la del capítulo X.

Antes de empezar este estudio, señalemos que las aplicaciones de la eliminación son numerosas: citemos la transformación de las ecuaciones algebraicas, el teorema de Bezout acerca de los sistemas de n hipersuperficies de $\mathcal{P}_n(K)$, los teoremas de «residuación» de Noether acerca de la intersección de dos curvas planas, y en general todos los problemas llamados de Geometría algebraica.

● En lo que sigue, K designa un cuerpo algebraicamente cerrado, de característica nula, por lo tanto infinito (cf. p. 108). Para la mayor parte de aplicaciones, $K = \mathbf{C}$.

§ VI.1 RESULTANTE DE DOS POLINOMIOS CON UNA SOLA VARIABLE

Sean dos polinomios con coeficientes en K :

$$(1) \quad \begin{aligned} f(X) &= a_0 X^m + a_1 X^{m-1} + \cdots + a_m & a_0 &\neq 0 \\ g(X) &= b_0 X^n + \cdots & & + b_n & b_0 &\neq 0. \end{aligned}$$

Buscamos las condiciones que ligan los a_i y los b_j para que las dos ecuaciones $f(x) = 0$, $g(x) = 0$ posean, por lo menos, una raíz común: ello equivale a decir que el polinomio $\text{mcd}(f, g)$ es de grado ≥ 1 . Ante todo establezcamos un lema:

VI.1.1 Sean dos polinomios $f(X) = \sum_{k=0}^m a_k X^{m-k}$, $g(X) = \sum_{k=0}^n b_k X^{n-k}$; para que
 || el mcd de f y g sea de grado ≥ 1 , es necesario y suficiente que existan
 || dos polinomios $U(X)$ y $V(X)$, tales que
 || (2) $Uf + Vg = 0$, $\text{gr}(U) \leq n - 1$, $\text{gr}(V) \leq m - 1$, $U \neq 0$, $V \neq 0$.

Demostración

a) Sea D el mcd de f y g ; hagamos $U = \frac{g}{D}$, $V = -\frac{f}{D}$. Si $\text{gr}(D) \geq 1$, los polinomios U y V satisfacen (2).

b) Recíprocamente, si U y V verifican (2), se tiene: $Uf = -Vg$. Descompongamos los dos miembros de esta relación en factores de grado 1 (es posible, puesto que K es algebraicamente cerrado). Como que $\text{gr}(U) \leq \text{gr}(g) - 1$, vemos que g contiene, a lo sumo, $n - 1$ factores que provienen de U ; luego, en la expresión de g , debe figurar, por lo menos, un factor de f . En otras palabras, el mcd de f y g es de grado ≥ 1 .

He aquí una demostración más formal que la precedente: si g y f son primos entre sí, g dividirá a U (T. de Gauss); la relación

$$\text{gr}(U) \leq \text{gr}(g) - 1$$

implicaría $U = 0$, lo cual es absurdo. c.q.d.

De VI.1.1 deduciremos las condiciones para que las ecuaciones (1) posean una raíz común. Para ello, escribimos las relaciones (2) escribiendo explícitamente U y V :

$$(3) \quad \begin{aligned} U &= \lambda_0 X^{n-1} + \lambda_1 X^{n-2} + \cdots + \lambda_{n-1}, & V &= \mu_0 X^{m-1} + \mu_1 X^{m-2} + \cdots + \mu_{m-1} \\ Uf + Vg &= \lambda_0(X^{n-1}f) + \lambda_1(X^{n-2}f) + \cdots + \lambda_{n-1}f + \\ &\quad + \mu_0(X^{m-1}g) + \mu_1(X^{m-2}g) + \cdots + \mu_{m-1}g. \end{aligned}$$

Puesto que $Uf + Vg = 0$, U y V son ambos no nulos si, y sólo si, uno de ellos es no nulo, es decir, si los coeficientes

$$\lambda_0, \lambda_1, \dots, \lambda_{n-1}, \mu_0, \mu_1, \dots, \mu_{m-1}$$

no son todos nulos. Según (3), la relación (2) puede, entonces, expresarse de la forma siguiente:

Los $m + n$ polinomios

$$(4) \quad f, Xf, \dots, X^{n-1}f, g, Xg, \dots, X^{m-1}g,$$

forman una familia ligada en el espacio vectorial $K[X]$. En la base canónica

$$(X^{m+n-1}, X^{m+n-2}, \dots, X^2, X, 1)$$

del espacio de los polinomios de grado $\leq m + n - 1$, la matriz de los polinomios (4) se escribe:

$$M = \begin{bmatrix} \overbrace{a_0 \ 0 \ \dots \ 0}^n & \overbrace{b_0 \ 0 \ \dots \ 0}^m \\ a_1 & b_1 & \ddots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\ a_m & \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & a_m & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & a_m & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & \dots & 0 & a_m & \vdots & \vdots \end{bmatrix} \quad (\text{matriz de Sylvester}).$$

Hemos demostrado, pues:

TEOREMA VI.1.2

Para que los dos polinomios

$$f(X) = \sum_{k=0}^m a_k X^{m-k}, \quad g(X) = \sum_{k=0}^n b_k X^{n-k}$$

tengan una raíz común, es necesario y suficiente que el determinante $R_{f,g}$, dado a continuación, sea nulo:

$$\bar{R}_{f,g} = \begin{vmatrix} a_0 & 0 & \dots & 0 & b_0 & 0 & \dots & 0 \\ a_1 & \ddots & & & \vdots & \ddots & & \\ \vdots & & \ddots & & \vdots & & \ddots & \\ a_m & & & a_0 & \vdots & & & b_0 \\ 0 & & & \vdots & b_n & & & \vdots \\ \vdots & & & \vdots & 0 & \ddots & & \\ 0 & & & a_m & 0 & \dots & & b_n \end{vmatrix} \quad (\text{determinante de Sylvester}).$$

DEFINICIÓN VI.1.1

Sean $u_0, u_1, \dots, u_m; v_0, v_1, \dots, v_n$ variables en el cuerpo de base K . El polinomio

$$(5) \quad R(u, v) = \det \begin{bmatrix} u_0 & 0 & \dots & 0 & v_0 & 0 & \dots & 0 \\ u_1 & \ddots & & & \vdots & \ddots & & \\ \vdots & & \ddots & & \vdots & & \ddots & \\ u_m & & & u_0 & \vdots & & & v_0 \\ 0 & & & \vdots & v_n & & & \vdots \\ \vdots & & & \vdots & 0 & \ddots & & \\ 0 & \dots & 0 & u_m & 0 & \dots & & v_n \end{bmatrix}$$

se llama **resultante** de los polinomios «generales»

$$f = \sum_{k=0}^m u_k X^{m-k} \quad \text{y} \quad g = \sum_{k=0}^n v_k X^{n-k}.$$

Vamos a dar ahora algunas propiedades de la resultante. En lo que sigue, si A es un subanillo cualquiera de K , designamos por $A[u, v]$ el anillo $A[u_0, \dots, u_m; v_0, \dots, v_n]$. Si k es el cuerpo de las fracciones de A , $k(u, v)$ designará el cuerpo de las fracciones de $A[u, v]$.

La fórmula (5) prueba que $R(u, v)$ contiene un término y sólo uno de la forma $u_0^n v_n^m$; este término es el único que contiene a u_0^n . Análogamente, $R(u, v)$ contiene un único término de la forma $(-1)^{mn} u_m^n v_0^m$, que es el único que contiene a u_m^n .

(5) prueba también que $R(u, v)$ es homogéneo de grado n respecto a u_0, u_1, \dots, u_m , y homogéneo de grado m respecto a v_0, v_1, \dots, v_n .

Pues si se multiplica cada u_i por $\lambda \in K$ (resp. cada v_j por λ), $R(u, v)$ queda evidentemente multiplicado por λ^n (resp. λ^m).

En particular, $R(u, v)$ es homogéneo de grado $m + n$ respecto al conjunto de las variables u, v .

TEOREMA VI.1.3

Si $u_0, u_1, \dots, u_m; v_0, \dots, v_n$ son variables en el cuerpo K , $R(u, v)$ es **irreducible** en el anillo $K[u, v]$ (en otras palabras, todo polinomio con coeficientes en K , que divida a $R(u, v)$ es o bien constante, o bien proporcional a $R(u, v)$).

Demostración (por recurrencia sobre el entero $m + n = \mu$). Si $\mu = 2$ (en cuyo caso $m = n = 1$) es fácil ver que $R(u, v) = u_0 v_1 - u_1 v_0$ es irreducible: en efecto, la forma cuadrática $u_0 v_1 - u_1 v_0$ es de rango 4 y no puede ser el producto de dos formas lineales de las 4 variables u_0, v_1, u_1, v_0 .

Supongamos, pues, verdadera la proposición para $m + n = \mu$ ($\mu \geq 2$) y demos-
tremos que es verdadera también para $m + n = \mu + 1$.

Si la suma $m + n = \mu + 1$ es ≥ 3 , uno de los enteros, por ejemplo, m para
fijar ideas, es ≥ 2 . Si hacemos $u_0 = 0$ en la relación $R(u, v) = 0$, es fácil ver que
la condición necesaria y suficiente para que los polinomios

$$u_1 X^{m-1} + u_2 X^{m-2} + \dots + u_m \quad \text{y} \quad v_0 X^n + v_1 X^{n-1} + \dots + v_n$$

tengan, por lo menos, una raíz común es que la relación obtenida sea $R(0, u_1, \dots, u_m; v_0, \dots, v_n) = 0$.

Precisando, si designamos por $R_0(u_1, \dots, u_m; v_0, \dots, v_n)$ la resultante de estos
polinomios, y si escribimos los determinantes R y R_0 , se verifica sin ninguna difi-
cultad la relación

$$R(0, u_1, \dots, u_m; v_0, \dots, v_n) = (-1)^n v_0 R_0(u_1, \dots, u_m; v_0, \dots, v_n).$$

Pero R_0 es, a menos de un cambio de notación, la resultante de dos polinomios
generales de grados respectivos $m - 1, n$, cuya suma de grados es μ . Por hipótesis
de recurrencia, R_0 es entonces irreducible, y el polinomio R no se podrá descom-
poner en un producto de polinomios $R = \Phi\Psi$, a menos que uno de los factores,
por ejemplo, Φ para fijar ideas, satisfaga

$$\Phi(0, u_1, \dots, u_m; v_0, \dots, v_n) = \alpha v_0,$$

con $\alpha = \text{Cte} \neq 0$. Por razones de homogeneidad, Φ sería de grado uno, por lo
tanto de la forma

$$\Phi(u_0, u_1, \dots, u_m; v_0, \dots, v_n) = \beta u_0 + \alpha v_0,$$

con $\alpha, \beta = \text{Cte}$; y la constante β sería también no nula ya que R no es divisible
por v_0 (contiene al término $u_0^u v_n^m$). Pero si $R(u, v)$ fuese divisible por $\beta u_0 + \alpha v_0$, con

$\alpha \neq 0$ y $\beta \neq 0$, tendríamos que, dos polinomios cualesquiera $f = u_0 X^m + \dots + u_m$, $g = v_0 X^n + \dots + v_n$ cuyos coeficientes u_0, v_0 verificaran $\alpha u_0 + \beta v_0 = 0$, poseerían siempre una raíz en común, lo cual es absurdo (pues multiplicando f y g por constantes convenientes, podemos limitarnos siempre al caso en que esta relación se verifica). Luego R es irreducible. c.q.d.

Nota. Para establecer que no se puede tener $P = \Phi\Psi$ con

$$\Phi = \alpha u_0 + \beta v_0,$$

podemos observar además que cada uno de los polinomios Φ, Ψ debería ser homogéneo respecto de las variables u_0, \dots, u_m y respecto de las variables v_0, \dots, v_n (pues el producto de dos polinomios homogéneos sólo puede ser homogéneo si ambos factores son homogéneos), lo que exige $\alpha = 0$ o $\beta = 0$.

Aplicación

En el capítulo XIV (p. 520) demostraremos la propiedad siguiente:

Si $P(X_1, \dots, X_n)$ es un polinomio irreducible con n variables (y coeficientes en K , algebraicamente cerrado y de característica nula), y si $F(X_1, \dots, X_n)$ es un polinomio con n variables, tal que las relaciones

$$(a_1, \dots, a_n) \in K^n \quad \text{y} \quad P(a_1, \dots, a_n) = 0$$

implican $F(a_1, \dots, a_n) = 0$, F es divisible por P .

Esta propiedad nos servirá para dar una expresión nueva de la resultante:

TEOREMA VI.1.4

Sean los polinomios con coeficientes en K

$$f(X) = \sum_{k=0}^m a_k X^{m-k}, \quad g(X) = \sum_{k=0}^n b_k X^{n-k}.$$

Designemos por $\alpha_1, \dots, \alpha_m$ las raíces de f , por β_1, \dots, β_n las de g . Se tiene

$$\begin{aligned} (6) \quad R_{f,g} &= a_0^n b_0^m \prod_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} (\alpha_i - \beta_j) = a_0^n \prod_{i=1}^m g(\alpha_i) \\ &= (-1)^{mn} b_0^m \prod_{i=1}^n f(\beta_i) = (-1)^{mn} R_{g,f}. \end{aligned}$$

Demostración. El segundo miembro de (6) es un polinomio simétrico respecto de $\alpha_1, \dots, \alpha_m$ por una parte, y respecto de β_1, \dots, β_n por otra parte. Según el teorema V.1.2, $P = \prod_{i,j} (\alpha_i - \beta_j)$ es un polinomio de grado n respecto de los a_k/a_0 ($1 \leq k \leq m$), y de grado m respecto de los b_k/b_0 ($1 \leq k \leq n$). Luego $a_0^n b_0^m P$ es un polinomio homogéneo de grado $m + n$ respecto de $(a_0, a_1, \dots, a_m; b_0, \dots, b_n)$. Demos a estas variables valores tales que $R(a_0, \dots, a_m; b_0, \dots, b_n) = 0$. Según el teorema VI.1.2, para estos valores, uno de los α_i será igual a uno de los β_j ; se tendrá pues: $a_0^n b_0^m P = 0$. Puesto que $R(u, v)$ es irreducible en $K[u, v]$, resulta de la propiedad recordada anteriormente que $a_0^n b_0^m P$ es divisible por $R_{f,g}$. Pero, dado que su grado es el mismo, le es proporcional ⁽¹⁾. Finalmente, utilizando la descomposición de f, g , se tiene:

$$P = \frac{1}{b_0^m} g(\alpha_1) \dots g(\alpha_m) = \frac{(-1)^{mn}}{a_0^n} f(\beta_1) \dots f(\beta_n),$$

lo que prueba que en P existe un único término de la forma

$$(\alpha_1 \dots \alpha_m)^n = (-1)^{mn} \frac{a_m^n}{a_0^n},$$

luego, en $a_0^n b_0^m P$, existe un único término de la forma $(-1)^{mn} b_0^m a_m^n$. Por lo tanto, el coeficiente de proporcionalidad entre $R(u, v)$ y $a_0^n b_0^m P$ es 1, de donde se sigue (6). c.q.d.

Aplicación

Designemos por $f_h(X)$ (resp. $g_h(X)$) al polinomio $f(X + h)$ (resp. $g(X + h)$). La fórmula (6) muestra también que se verifica:

$$R_{f,g} = R_{f_h, g_h},$$

pues las diferencias $\alpha_i - \beta_j$ son invariantes, así como los coeficientes a_0 y b_0 (invariancia de la resultante por traslación). En ciertos casos, se podrá utilizar esta propiedad para simplificar los cálculos.

He aquí otra propiedad importante de la resultante:

⁽¹⁾ De hecho, el razonamiento no es completo. En rigor, se tiene $a_0^n b_0^m P = 0$ para todos los sistemas $(a_0, \dots, a_m; b_0, \dots, b_n)$ tales que $R(a_0, \dots, a_m; b_0, \dots, b_n) = 0$ con $a_0 b_0 \neq 0$; pero en el capítulo XIV veremos que el teorema de los ceros de Hilbert se aplica también en este caso, pues ni a_0 ni b_0 aparecen en R como factor.

TEOREMA VI.1.5

Sea $R(u, v)$ la resultante de los polinomios generales

$$f = \sum_{k=0}^m u_k X^{m-k}, \quad g = \sum_{k=0}^n v_k X^{n-k}.$$

Existen polinomios $A, B \in \mathbf{Z}[u, v; X]$, de grados respectivos $\leq n-1$ y $\leq m-1$ en X , determinados de forma única, tales que se cumple:

$$(7) \quad R(u, v) = A(u, v, X)f(X) + B(u, v, X)g(X).$$

Demostración

a) *Existencia.* Consideremos la matriz de Sylvester de f y g , llamémosla M , y sean L_1, L_2, \dots, L_{m+n} sus vectores-fila:

$$M = \begin{bmatrix} u_0 & 0 & \dots & 0 & v_0 & 0 & \dots & 0 \\ u_1 & u_0 & & & v_1 & & & \\ \vdots & \vdots & \ddots & & \vdots & & & \\ u_m & & & u_0 & v_{n-1} & & & v_0 \\ 0 & u_m & & & v_n & & & \\ \vdots & \vdots & \ddots & & \vdots & & & \\ 0 & \dots & & u_m & 0 & \dots & 0 & v_n \end{bmatrix}$$

Hagamos:

$$(8) \quad \begin{aligned} L &= X^{m+n-1} L_1 + X^{m+n-2} L_2 + \dots + L_{m+n} \\ &= [X^{n-1} f(X), X^{n-2} f(X), \dots, 1 \cdot f(X), X^{m-1} g(X), \\ &\quad X^{m-2} g(X), \dots, 1 \cdot g(X)]. \end{aligned}$$

Sea N la matriz cuyas filas son $L_1, L_2, \dots, L_{m+n-1}, L$. Según (8) se tiene, en primer lugar:

$$\det(N) = \det(M) = R(u, v).$$

Por otro lado, desarrollando directamente $\det(N)$ según los elementos de L , se obtiene una relación de la forma (7).

b) *Unicidad.* Si existieran dos parejas distintas de polinomios de grados $\leq n - 1$ y $\leq m - 1$ en X que verifican (7), se deduciría, por diferencia, la existencia de polinomios no nulos A, B de grados $\leq n - 1$ y $\leq m - 1$ en X , tales que

$$Af(X) + Bg(X) = 0.$$

Pero tal relación es imposible, pues al ser $(u), (v)$ variables, los polinomios f y g son primos entre sí en $\mathbf{Q}[u, v, X]$ (para verlo basta con dar a (u) y a (v) valores de K tales que los polinomios f y g no posean ninguna raíz común en X).]]

§ VI.2 ALGUNOS EJEMPLOS DE CÁLCULO DE RESULTANTES

- 1) Dos polinomios de grado 1: $f = a_1 X + a_2, g = b_1 X + b_2$;

$$R(f, g) = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1.$$

- 2) Dos polinomios de grado 2:

$$f = a_0 X^2 + a_1 X + a_2, g = b_0 X^2 + b_1 X + b_2;$$

$$R_{f,g} = \begin{vmatrix} a_0 & 0 & b_0 & 0 \\ a_1 & a_0 & b_1 & b_0 \\ a_2 & a_1 & b_2 & b_1 \\ 0 & a_2 & 0 & b_2 \end{vmatrix}.$$

Tras el desarrollo, se obtiene:

$$R_{f,g} = \begin{vmatrix} a_0 & a_2 \\ b_0 & b_2 \end{vmatrix}^2 - \begin{vmatrix} a_0 & a_1 \\ b_0 & b_1 \end{vmatrix} \cdot \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Aplicación

El cuerpo de base es \mathbf{R} , y $a_0 = b_0 = 1$. Si α_1, α_2 (resp. β_1, β_2) designan las raíces de f (resp. de g), sabemos, según VI.1.4, que

$$R = (\alpha_1 - \beta_1)(\alpha_1 - \beta_2)(\alpha_2 - \beta_1)(\alpha_2 - \beta_2).$$

Supongamos que α_1 y α_2 son *reales*; si β_1 y β_2 son complejos conjugados,

$$R = |\alpha_1 - \beta_1|^2 |\alpha_2 - \beta_1|^2 > 0.$$

Luego si $R < 0$, β_1 y β_2 son reales, y en este caso, la fórmula anterior prueba que sólo tres de los términos $(\alpha_i - \beta_j)$ tienen el mismo signo, por lo tanto que se tiene (si suponemos $\alpha_1 < \alpha_2$ y $\beta_1 < \beta_2$), o bien $\alpha_1 < \beta_1 < \alpha_2 < \beta_2$, o bien $\beta_1 < \alpha_1 < \beta_2 < \alpha_2$. En resumen, si $R < 0$, y si una de las ecuaciones $f = 0$, $g = 0$ tiene sus raíces reales, la otra también tiene sus raíces reales, y ambos pares de raíces están intercalados.

3) La ecuación $f(X) = 0$ (f , polinomio) posee una raíz múltiple si, y sólo si, $R_{f,f'} = 0$. Por ejemplo, si $f = X^3 + pX + q$, $f' = 3X^2 + p$,

$$R_{f,f'} = \begin{vmatrix} 1 & 0 & 3 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ p & 0 & p & 0 & 3 \\ q & p & 0 & p & 0 \\ 0 & q & 0 & 0 & p \end{vmatrix} = 4p^3 + 27q^2.$$

4) Calculemos $R_{f,g}$ cuando

$$f(X) = X^n - 1 \quad \text{y} \quad g(X) = a_{n-1}X^{n-1} + \dots + a_1X + a_0$$

(cuerpo de base: \mathbf{C}).

En virtud del teorema VI.1.4, tenemos por una parte:

$$(1) \quad R(f, g) = \prod_{\xi \in U_n} g(\xi),$$

en donde U_n designa el conjunto de las raíces n -ésimas de 1.

Por otra parte, el método de Sylvester nos conduce a la fórmula:

$$R_{f,g} = \begin{vmatrix} 1 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 & -1 \\ a_{n-1} & \dots & a_1 & a_0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & a_1 & a_0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & a_{n-1} & a_{n-1} & \dots & a_0 & 0 \end{vmatrix}$$

En este determinante, sumamos la k -ésima columna a la $(n + k)$ -ésima ($1 \leq k \leq n - 1$): se obtiene $R(f, g)$ bajo la forma de un determinante de orden n :

$$(2) \quad R_{f,g} = \begin{vmatrix} a_0 & a_{n-1} & a_{n-2} & \cdots & a_1 \\ a_1 & a_0 & & & a_2 \\ a_2 & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{vmatrix},$$

(determinante circulante de a_0, a_1, \dots, a_{n-1}), que se designa por

$$C(a_0, a_1, \dots, a_{n-1}).$$

Teniendo en cuenta el signo de a_0^n , (1) y (2) proporcionan más rápidamente que el método del capítulo X, § 2 la identidad:

$$C(a_0, a_1, \dots, a_{n-1}) = \prod_{\zeta \in U_n} (a_{n-1} \zeta^{n-1} + a_{n-2} \zeta^{n-2} + \cdots + a_1 \zeta + a_0),$$

Método de Cayley

El cálculo de la resultante es, en general, muy difícil, pues el determinante de Sylvester es de orden elevado. Por ejemplo, para dos ecuaciones de grado 3, es de orden 6. Damos ahora un método, debido a Cayley, que requiere un determinante de orden menos elevado.

Supongamos que f y g tienen el mismo orden m :

$$f(X) = \sum_{k=0}^m a_k X^{m-k}, \quad g(X) = \sum_{k=0}^m b_k X^{m-k}.$$

Si las ecuaciones $f = 0$ y $g = 0$ tienen una raíz común x_0 , el polinomio h , dado por:

$$h(Y) = f(x_0) g(Y) - g(x_0) f(Y),$$

es nulo. Pero h siempre es divisible por $Y - x_0$. Luego el polinomio:

$$Q(x_0, Y) = \frac{h(Y)}{Y - x_0} \quad \text{debe ser nulo.}$$

Ordenemos $Q(x_0, Y)$ según las potencias de Y :

$$h(Y) = \sum_{k=0}^m (f(x_0) b_k - g(x_0) a_k) Y^{m-k} = \sum_{\substack{0 \leq k \leq m \\ 0 \leq l \leq m}} (b_k a_l - a_k b_l) x_0^{m-l} Y^{m-k},$$

$$(3) \quad h(Y) = \sum_{0 \leq k < l \leq m} (b_k a_l - a_k b_l) x_0^{m-l} Y^{m-l} (Y^{l-k} - x_0^{l-k}).$$

Si llevamos a (3) las relaciones:

$$Y^{l-k} - x_0^{l-k} = (Y - x_0) (Y^{l-k-1} + Y^{l-k-2} x_0 + \dots + x_0^{l-k-1})$$

y tenemos en cuenta que $Q(x_0, Y) = \frac{h(Y)}{Y - x_0}$, se obtiene:

$$Q(x_0, Y) = \sum_{\substack{0 \leq k < l \leq m \\ 0 \leq q \leq l-k-1}} (b_k a_l - a_k b_l) Y^{m-l+l-k-q-1} x_0^{m-l+q},$$

que haciendo $\lambda = k + q$ en los exponentes de dicha fórmula, da

$$(3) \quad Q(x_0, Y) = \sum_{\substack{0 \leq k, l \leq m \\ k < l, 0 \leq \lambda \leq l-1}} (b_k a_l - a_k b_l) Y^{m-\lambda-1} x_0^{m+\lambda-k-l}$$

$$= \sum_{\lambda=0}^{m-1} A_\lambda(x_0) Y^{m-\lambda-1},$$

con

$$(4) \quad A_\lambda(x_0) = \sum_{\substack{0 \leq k < l \leq m \\ l \geq \lambda+1}} (b_k a_l - a_k b_l) x_0^{m+\lambda-k-l}.$$

Puesto que el polinomio $Q(x_0, Y)$ es nulo, todos sus coeficientes $A_\lambda(x_0)$ son nulos. $A_\lambda(x_0)$ es un polinomio de grado $\leq m - 1$ en x_0 . Si hacemos

$$A_\lambda(x_0) = \sum_{0 \leq j \leq m-1} \alpha_{j\lambda} x_0^j,$$

el sistema de ecuaciones lineales y homogéneas:

$$(5) \quad \sum_{0 \leq j \leq m-1} \alpha_{j\lambda} u_j = 0 \quad (0 \leq \lambda \leq m-1)$$

admite una solución no trivial $(1, x_0, x_0^2, \dots, x_0^{m-1})$. Luego el determinante $\det [a_{j\lambda}]$ de (5) debe ser nulo. Pero según (3), $\det [a_{j\lambda}]$ es de grado $2m$ respecto de $a_0, \dots, a_m; b_0, \dots, b_m$. Puesto que el $\det [a_{j\lambda}]$ es nulo para todos los valores de las variables que anulan a $R_{f,g}$, vemos que $\det [a_{j\lambda}] = CR_{f,g}$, en donde C es una constante $\neq 0$. Se obtiene, de esta manera, un determinante de orden m . Podemos enunciar:

REGLA PRÁCTICA

Para hallar (a menos de un factor de proporcionalidad) la resultante de dos polinomios f y g del mismo grado m , es suficiente calcular el determinante $\det [a_{j\lambda}]$, en donde $a_{j\lambda}$ es el coeficiente de x_0^j en el término en $Y^{m-\lambda-1}$ del polinomio:

$$Q(x_0, Y) = [f(x_0)g(Y) - g(x_0)f(Y)](Y - x_0).$$

Como ejemplo, calculemos $R(f, g)$ cuando

$$f(X) = a_0 X^3 + a_1 X^2 + a_2 X + a_3,$$

$$g(X) = b_0 X^3 + b_1 X^2 + b_2 X + b_3.$$

Se obtiene:

$$\begin{aligned} f(x)g(Y) - g(x)f(Y) &= (a_1 b_0 - a_0 b_1) x^2 Y^2 (Y - x) + \\ &+ (a_2 b_0 - a_0 b_2) x Y (Y^2 - x^2) + (a_3 b_0 - a_0 b_3) (Y^3 - x^3) \\ &+ (a_2 b_1 - a_1 b_2) x Y (Y - x) + (a_3 b_2 - a_2 b_3) (Y - x) \\ &+ (a_3 b_1 - a_1 b_3) (Y^2 - x^2). \end{aligned}$$

$$\begin{aligned} Q(x_0, Y) &= [(a_1 b_0 - a_0 b_1) x^2 + (a_2 b_0 - a_0 b_2) x + a_3 b_0 - a_0 b_3] Y^2 + \\ &+ [(a_2 b_0 - a_0 b_2) x^2 + \{ (a_3 b_0 - a_0 b_3) + (a_2 b_1 - a_1 b_2) \} x \\ &+ a_3 b_1 - a_1 b_3] Y + (a_3 b_0 - a_0 b_3) x^2 + [a_3 b_1 - a_1 b_3] x \\ &+ a_3 b_2 - a_2 b_3. \end{aligned}$$

Se deduce la resultante:

$$(5) \quad R(f, g) = \begin{vmatrix} a_1 b_0 - a_0 b_1 & a_2 b_0 - a_0 b_2 & a_3 b_0 - a_0 b_3 \\ a_2 b_0 - a_0 b_2 & (a_3 b_0 - a_0 b_3) + (a_2 b_1 - a_1 b_2) & a_3 b_1 - a_1 b_3 \\ a_3 b_0 - a_0 b_3 & a_3 b_1 - a_1 b_3 & a_3 b_2 - a_2 b_3 \end{vmatrix}.$$

Se observará que, de forma general, la matriz $[a_{j\lambda}]$ del sistema (4) es simétrica.

Nota. Si f y g no poseen el mismo grado, el método de Cayley se aplica tomando ciertas precauciones; supongamos que

$$f(X) = a_0 X^m + \cdots + a_m, \quad g(X) = b_0 X^n + \cdots + b_n,$$

con $n < m$, $a_0 \neq 0$ y $a_m \neq 0$; $x = 0$ no es una raíz común de f y g . Luego para que f y g posean una raíz común, es necesario y suficiente que f y $X^{m-n}g$ tengan una raíz común, esto es,

$$R_{f, X^{m-n}g} = 0.$$

Pero $R_{f, X^{m-n}g}$ es de grado $2m$, mientras que $R_{f, g}$ es de grado $m+n$. $R_{f, X^{m-n}g}$ es, pues, el producto de $R_{f, g}$ por un factor de grado $m-n$. La fórmula (5) del § 1 prueba inmediatamente que este factor es a_m^{m-n} . De hecho, en la expresión de $R_{f, X^{m-n}g}$ obtenida por el método de Cayley, a_m aparecerá como factor en las $m-n$ últimas filas o columnas del determinante obtenido.

Método de reducción del grado

Los ejemplos nos permitirán entender suficientemente este método. Como se verá, permite obtener rápidamente una condición *necesaria* para que dos ecuaciones posean una raíz en común. Pero conduce a *factores parásitos* (la ecuación es, en general, de grado más elevado que la resultante) lo cual obliga a *comprobar* los cálculos a fin de asegurar las *condiciones suficientes*.

Este método consiste en calcular el mcd por medio del algoritmo de Euclides.

Ejemplo 1

Consideremos

$$f(X) = a_0 X^3 + a_1 X^2 + a_2 X + a_3, \quad a_0 a_3 \neq 0, \quad g(X) = b_0 X^2 + b_1 X + b_2.$$

Si f y g poseen una raíz en común, también la poseen $f_1 = b_0 f - a_0 Xg$ y $g_1 = g$:

$$\begin{aligned} f_1 &= (a_1 b_0 - a_0 b_1) X^2 + (b_0 a_2 - a_0 b_2) X + b_0 a_3, \\ g_1 &= b_0 X^2 + b_1 X + b_2. \end{aligned}$$

Análogamente, $b_0 f_1 - (a_1 b_0 - a_0 b_1) g_1 = f_2$, y $g_2 = g$, poseen una raíz común y se tiene:

$$f_2 = AX + B, \quad \text{con } A = b_0(b_0 a_2 - a_0 b_2) - b_1(a_1 b_0 - a_0 b_1), \\ B = b_0^2 a_3 - b_2(a_1 b_0 - a_0 b_1),$$

finalmente, $f_3 = f_2$ y $g_3 = b_0 X f_2 - A g_2$ tienen una raíz común. Luego se tiene:

$$g_3 = (b_0 B - A b_1) X - A b_2.$$

Debemos tener, pues:

$$(6) \quad \begin{vmatrix} A & B \\ b_0 B - A b_1 & -A b_2 \end{vmatrix} = 0.$$

Según la teoría, el determinante Δ del primer miembro de (6) es un múltiplo de $R(f, g)$. Sin embargo, vemos que Δ es de grado 7 respecto del conjunto de los a_i y de los b_j , mientras que $R(f, g)$ es de grado 5. Δ contiene, pues, un factor parásito de grado 2.

Si terminamos los cálculos, obtenemos:

$$\Delta = \begin{vmatrix} b_0^2 a_2 - a_0 b_0 b_2 - a_1 b_0 b_1 + a_0 b_1^2, & -b_0^2 a_3 - a_1 b_0 b_2 + a_0 b_1 b_2 \\ b_0^3 a_3 - b_0^2 a_1 b_2 + 2 a_0 b_0 b_1 b_2 - b_0^2 a_2 b_1 + a_1 b_0 b_1^2 - a_0 b_1^3, \\ & -a_2 b_0^2 b_2 + a_0 b_0 b_2^2 + a_1 b_0 b_1 b_2 - a_0 b_1^2 b_2 \end{vmatrix}$$

$R(f, g)$ contiene un solo término en $a_0^2 b_2^3$. En Δ , el coeficiente de $a_0^2 b_2^3$ es $-b_0^2 a_0^2 b_2^3$. Luego el factor parásito es b_0^2 .

Ejemplo 2

Eliminar x entre las ecuaciones:

$$\begin{cases} x^4 + x + a = 0, \\ ax^4 + x^3 + 1 = 0. \end{cases}$$

El método de reducción del grado conduce, suponiendo que $a \neq 0$, a la ecuación de primer grado:

$$(a^3 - a^2 - a)x - (a^3 - a^2 - a) = 0.$$

Si $a^3 - a^2 - a \neq 0$, la raíz común a ambas ecuaciones es, pues, $x = 1$, de donde se obtiene inmediatamente $a = -2$.

Si $a^3 - a^2 - a = 0$, las dos ecuaciones tienen en común *dos* raíces. De manera más precisa, el mcd de los primeros miembros es entonces de grado 2.

Si $a = 0$ las dos ecuaciones tienen *tres* raíces comunes, las raíces cúbicas de -1 . Es interesante calcular directamente la resultante por el método de Cayley. Haciendo:

$$\begin{aligned} f(x) &= x^4 + x + a \\ g(x) &= ax^4 + x^3 + 1, \end{aligned}$$

se tiene

$$\frac{f(y)g(x) - f(x)g(y)}{x - y} = Ay^3 + By^2 + Cy + D,$$

con

$$\begin{aligned} A &= -x^3 + ax + a^2 - 1, & B &= ax^2 + a^2x + a, \\ C &= ax^3 + a^2x^2 + ax, & D &= (a^2 - 1)x^3 + ax^2 - 1. \end{aligned}$$

Se obtiene pues:

$$R_{f,g} = \begin{vmatrix} -1 & 0 & a & a^2 - 1 \\ 0 & a & a^2 & a \\ a & a^2 & a & 0 \\ a^2 - 1 & a & 0 & -1 \end{vmatrix}.$$

Un cálculo fácil prueba:

$$R_{f,g} = a^3(a + 2)(a^2 - a - 1)^2.$$

§ VI.3 APLICACIÓN DE LA ELIMINACIÓN A LA TRANSFORMACIÓN DE TSCHIRNHAUS

Dada $f(x) = 0$ (en donde $f(x) = \sum_{k=0}^m a_k x^{m-k}$) una ecuación algebraica con coeficientes en K , sea $f(x) = \prod_{k=1}^m (X - \alpha_k)$. Se considera $\varphi(x) = \frac{P(x)}{Q(x)}$ una fracción racional, cuyo denominador $Q(x)$ no se anula para $x = \alpha_1, \dots, x = \alpha_m$, lo que equivale a decir que f y Q son primos entre sí. Por definición, la transformada de f por φ es la ecuación cuyas raíces son

$$\varphi(\alpha_1), \varphi(\alpha_2), \dots, \varphi(\alpha_m).$$

Con más precisión, la transformada es la ecuación $g(x) = 0$, donde $g(x)$ es el polinomio $\prod_{k=1}^m (X - \varphi(\alpha_k))$.

Vamos a ver que siempre es posible reducirse al caso en que φ es un polinomio.

Según el teorema de Bezout, existen polinomios $U(x)$ y $V(x)$ tales que

$$Uf + VQ = 1,$$

de donde

$$\varphi = \frac{PUf}{Q} + \frac{PVQ}{Q} = \frac{PUf}{Q} + PV.$$

Haciendo $X = \alpha_i$ ($1 \leq i \leq m$) en esta relación, se obtiene (puesto que $f(\alpha_i) = 0$), $\varphi(\alpha_i) = P(\alpha_i) V(\alpha_i)$, lo que demuestra que la transformada de f por φ es igual a su transformada por $\psi = PV$.

DEFINICIÓN VI.3.1

Sea $f(x) = 0$ una ecuación algebraica, en donde $f(x) = \sum_{k=0}^m a_k x^{m-k}$. La transformada de esta ecuación por la **transformación de Tschirnhaus** asociada al polinomio $\varphi(x) = \sum_{k=0}^n b_k x^{n-k}$ es la ecuación $g(x) = 0$, con $g(x) = \prod_{k=1}^m (X - \varphi(\alpha_k))$ y $f(x) = \prod_{k=1}^m (X - \alpha_k)$. El grado de φ es el grado de la transformación.

Nota. Podemos reducirnos siempre a una transformación de grado $\leq m - 1$. En efecto, la división euclídea de φ por f nos da:

$$\varphi = fh + R, \quad \text{gr}(R) \leq m - 1.$$

Si en esta relación hacemos $x = \alpha_i$, ($1 \leq i \leq m$), se obtiene:

$$\varphi(\alpha_i) = R(\alpha_i),$$

lo que prueba que la transformada de f por φ es igual a la de f por R .

Pretendemos calcular la transformada de f por el polinomio φ . Un primer método consiste en aplicar la definición:

$$(1) \quad (y - \varphi(\alpha_1))(y - \varphi(\alpha_2)) \dots (y - \varphi(\alpha_m)) = 0.$$

Desarrollando (1), se obtiene una ecuación cuyos coeficientes son *funciones simétricas* en $\alpha_1, \alpha_2, \dots, \alpha_m$; podemos pues, en principio, calcularlas en función de a_0, a_1, \dots, a_m . En el capítulo V hemos dado algunos ejemplos de este método.

Un segundo método es la *utilización de la eliminación*: el primer miembro de (1) es, salvo un factor multiplicativo, la resultante de los dos polinomios:

$$\begin{cases} a_0 x^m + \dots + a_m \\ \varphi(x) - y \end{cases}$$

Haciendo: $g_y(x) = \varphi(x) - y$, la condición buscada se escribe entonces:

$$(2) \quad R_{f, g_y} = 0.$$

R_{f, g_y} es de grado m en y (§ 1), siendo $(-1)^m a_0^n y^m$ el término en y^m . El método de Cayley será muy ventajoso, si bien para $n < m$ introduce un factor parásito a_m^{m-n} , pero en la práctica, este hecho no es molesto.

Ejemplos

1) Transformemos

$$(3) \quad x^3 + px + q = 0 \quad \text{por} \quad (4) \quad y = b_0 x^2 + b_1 x + b_2.$$

El método de Cayley proporciona la transformada en la forma siguiente (tras simplificar la última fila por q):

$$\begin{vmatrix} -b_1 & pb_0 - (b_2 - y) & qb_0 \\ pb_0 - (b_2 - y) & qb_0 + pb_1 & qb_1 \\ b_0 & b_1 & b_2 - y \end{vmatrix} = 0.$$

Desarrollando, se obtiene:

$$\begin{aligned} y^3 + [2pb_0 - 3b_2]y^2 + [(pb_0 - b_2)(pb_0 - 3b_2) + b_1(pb_1 + 3qb_0)]y \\ + qb_1^3 + (pb_0 - b_2)(2qb_0b_1 - b_2\{pb_0 - b_2\}) \\ - (qb_0 + pb_1)(b_1b_2 + qb_0^2) = 0. \end{aligned}$$

Impongamos: $b_0 = 3$. Si escribimos que el término en y^2 de esta transformación es nulo, obtenemos: $b_2 = 2p$. Escribamos entonces que el término en y es nulo, y b_1 queda determinado por la ecuación de *segundo* grado:

$$pb_1^2 + 9qb_1 - 3p^2 = 0, \quad \text{cuyas raíces son} \quad \frac{-9q \pm \sqrt{3}\sqrt{4p^3 + 27q^2}}{2p}.$$

Por lo tanto, mediante una transformación (4) de la que se sepa obtener en forma explícita los coeficientes, es posible transformar (3) en una ecuación de la forma: $y^3 - B = 0$. Puesto que la resolución de esta última ecuación es inmediata, se ha hallado así un nuevo método de resolución de (3) — teniendo en cuenta los resultados del § 5.

2) Transformemos

$$(5) \quad x^4 + ax^2 + bx + c = 0 \quad \text{por} \quad (6) \quad y = x^2 + b_1 x + b_2.$$

El método de Cayley conduce a la transformada:

$$\begin{vmatrix} b_1 & b_2 - y - a & -b & -1 \\ b_2 - y - a & -b - ab_1 & -c - bb_1 & -b_1 \\ -b & -c - bb_1 & -cb_1 - b(b_2 - y) & -(b_2 - y) \\ -1 & -b_1 & -(b_2 - y) & 0 \end{vmatrix} = 0.$$

La descomposición de este determinante Δ con la ayuda de sus menores de orden 2:

$$\begin{aligned} \Delta = & \begin{vmatrix} b_2 - y - a & -b \\ -b - ab_1 & -c - bb_1 \end{vmatrix} \cdot \begin{vmatrix} -b & -(b_2 - y) \\ -1 & 0 \end{vmatrix} + \\ & + \begin{vmatrix} b_1 & b \\ b_2 - y - a & c + bb_1 \end{vmatrix} \cdot \begin{vmatrix} -c - bb_1 & -(b_2 - y) \\ -b_1 & 0 \end{vmatrix} \\ & + \begin{vmatrix} b_1 & b_2 - y - a \\ b_2 - y - a & -b - ab_1 \end{vmatrix} \cdot \begin{vmatrix} -cb_1 - b(b_2 - y) & -(b_2 - y) \\ -(b_2 - y) & 0 \end{vmatrix} \\ & + \begin{vmatrix} b_1 & -1 \\ b_2 - y - a & -b_1 \end{vmatrix} \cdot \begin{vmatrix} -c - bb_1 & -cb_1 - b(b_2 - y) \\ -b_1 & -(b_2 - y) \end{vmatrix} \\ & + \begin{vmatrix} b_2 - y - a & -1 \\ -b - ab_1 & -b_1 \end{vmatrix} \cdot \begin{vmatrix} b & -cb_1 - b(b_2 - y) \\ 1 & -(b_2 - y) \end{vmatrix} \\ & + \begin{vmatrix} b & 1 \\ c + bb_1 & b_1 \end{vmatrix} \cdot \begin{vmatrix} b & c + bb_1 \\ 1 & b_1 \end{vmatrix}, \end{aligned}$$

nos da el coeficiente de y^3 , a saber B_1 , y el coeficiente B_3 de y , en la transformada:

$$B_1 = 2a - 4b_2, \quad B_3 = \psi(a, b, c; b_1, b_2),$$

en donde ψ es de grado 3 respecto de b_1 (el término en b_1^3 de ψ es: $b \cdot b_1^3$). Tomando $b_2 = \frac{a}{2}$, se tendrá $B_1 = 0$. La ecuación $B_3 = 0$ determinará entonces b_1 por medio de una ecuación de grado 3, que sabemos resolver. En otras palabras, *es posible, mediante una transformación de la forma (6), transformar la ecuación (5) en una ecuación de la forma: $y^4 + B_2 y^2 + B_4 = 0$* . Puesto que esta última ecuación se resuelve de forma elemental, hemos transformado la ecuación (5) en una ecuación de tercer grado — teniendo en cuenta los resultados del § 5.

Es posible demostrar que toda ecuación de grado m ,

$$f(x) \equiv a_0 x^m + a_1 x^{m-1} + \dots + a_m = 0,$$

se puede transformar en una ecuación: $b_0 y^m + b_1 y^{m-1} + \dots + b_m = 0$, en que $b_1 = b_2 = b_3 = 0$, con la ayuda de una transformación de grado 4: $\varphi(x) = \sum_{i=0}^4 a_i x^i$, cuyos coeficientes a_i vienen determinados por ecuaciones de grado 2 ó 3.

3) Transformemos la ecuación $x^n - 1 = 0$ por el polinomio:

$$y = a_0 + a_1 x + \dots + a_{n-1} x^{n-1}.$$

Equivale a eliminar x entre las ecuaciones $f(x) = 0$, $g(x) = 0$, con:

$$f(x) = x^n - 1,$$

$$g(x) = a_{n-1} x^{n-1} + \dots + a_1 x + a_0 - y.$$

Suponemos que el cuerpo de base es \mathbf{C} .

Según el ejemplo 4) del § 2, la ecuación transformada se escribe $R = 0$, con

$$(7) \quad R = R_{f,g} = \begin{vmatrix} a_0 - y & a_{n-1} & \dots & a_1 \\ a_1 & a_0 - y & a_{n-1} & \dots & a_2 \\ \vdots & & & & \vdots \\ a_{n-1} & \dots & a_1 & a_0 - y \end{vmatrix} \\ = \prod_{\zeta \in U_n} (a_{n-1} \zeta^{n-1} + a_{n-2} \zeta^{n+2} + \dots + a_1 \zeta + a_0 - y).$$

4) Busquemos la transformada de la ecuación $f(x) = 0$, en donde

$$f(x) = a_0 x^m + a_1 x^{m-1} + \dots + a_m,$$

por medio del polinomio $y = x^k$ ($k \in \mathbf{N}$).

Haciendo $g_y(x) = x^k - y$, debemos calcular $R(f, g_y)$. A este fin, designamos por U_k el grupo de las raíces k -ésimas de 1 en \mathbf{C}^* .

Vamos a dar un significado correcto a la idea que consiste en hacer $y = z^k$, lo cual es perfectamente natural, pues se intenta explotar la fórmula:

$$x^k - z^k = \prod_{\zeta \in U_k} (x - \zeta z)$$

Descomponemos $f(x)$ en factores lineales:

$$f(x) = a_0 \prod_{j=1}^n (x - x_j) ;$$

formamos $P(y) = \prod_{\zeta \in U_k} f(\zeta y)$:

$$\begin{aligned} P(y) &= a_0^k \prod_{\substack{1 \leq j \leq n \\ \zeta \in U_k}} (\zeta y - x_j) = a_0^k \prod_{1 \leq j \leq n} (-1)^{k+1} (y^k - x_j^k) \\ &= (-1)^{n(k+1)} a_0^k \prod_{1 \leq j \leq n} (y^k - x_j^k) . \end{aligned}$$

Puesto que la transformada de la ecuación propuesta es $\prod_{1 \leq j \leq n} (z - x_j^k) = 0$, vemos que se obtiene haciendo $z = y^k$ en la expresión:

$$P(y) = \prod_{\zeta \in U_k} f(\zeta y) .$$

— Para ilustrar este procedimiento, calculemos la transformada de

$$f(x) = x^5 + x^3 + x^2 + 2x + 3 = 0 \quad \text{por } y = x^2 .$$

Debemos hacer $y = x^2$ en $f(x)f(-x)$, pero, $f(x) = P(x^2) + xQ(x^2)$, con

$$P(y) = y + 3, \quad Q(y) = y^2 + y + 2 ;$$

de donde

$$f(x)f(-x) = P^2(y) - yQ^2(y) ,$$

lo cual nos conduce a la transformada:

$$y^5 + 2y^4 + 5y^3 + 3y^2 - 2y - g = 0 .$$

(El método aquí empleado se extiende a toda transformación por $y = x^2$.)

— Busquemos ahora la transformada de

$$f(x) = x^m + a_1 x^{m-1} + \dots + a_m = 0 \quad \text{por } y = x^3 .$$

Si $j = e^{2\pi i/3}$, debemos hacer $y = x^3$ en el producto:

$$\Phi = f(x)f(jx)f(j^2 x).$$

Pero podemos escribir:

$$f(x) = P(x^3) + xQ(x^3) + x^2 R(x^3),$$

en donde P, Q, R son polinomios.

Se tiene, pues:

$$\begin{aligned} \Phi &= (P(x^3) + xQ(x^3) + x^2 R(x^3)) (P(x^3) + jxQ(x^3) + j^2 x^2 R(x^3)) \\ &\quad (P(x^3) + j^2 xQ(x^3) + jx^2 R(x^3)) = P^3 + x^3 Q^3 + x^6 R^3 - 3x^3 PQR. \end{aligned}$$

La transformada buscada es, entonces:

$$P^3(y) + yQ^3(y) + y^2 R^3(y) - 3yP(y)Q(y)R(y) = 0.$$

— De forma análoga, el lector demostrará que la transformada de

$$f(x) = x^m + a_1 x^{m-1} + \dots + a_m \quad \text{por } y = x^k$$

se obtiene de la manera siguiente: se escribe $f(x)$ en la forma:

$$f(x) = P_0(x^k) + xP_1(x^k) + \dots + x^{k-1} P_{k-1}(x^k).$$

Según la fórmula (7) del ejemplo 3), la transformada se puede poner en forma de determinante:

$$(8) \quad \begin{vmatrix} P_0(y) & xP_1(y) & x^2 P_2(y) & \dots & x^{k-1} P_{k-1}(y) \\ x^{k-1} P_{k-1}(y) & P_0(y) & xP_1(y) & \dots & x^{k-2} P_{k-2}(y) \\ \vdots & & & & \vdots \\ xP_1(y) & x^2 P_2(y) & \dots & \dots & P_0(y) \end{vmatrix} = 0.$$

A fin de obtener el resultado en forma de un determinante que dependa sólo de y , en (8) se multiplica la segunda fila por x , la tercera por x^2 , etc., luego, en el determinante obtenido, la segunda columna por x^{k-1} , la tercera por x^{k-2} , etc. Después de simplificar las $n - 1$ últimas filas por y , se obtiene:

$$\begin{vmatrix} P_0(y) & yP_1(y) & yP_2(y) & \dots & yP_{k-1}(y) \\ P_{k-1}(y) & P_0(y) & P_1(y) & \dots & P_{k-2}(y) \\ P_{k-2}(y) & yP_{k-1}(y) & P_0(y) & P_1(y) & \dots \\ \vdots & & & & \vdots \\ P_1(y) & yP_2(y) & \dots & yP_{k-1}(y) & P_0(y) \end{vmatrix} = 0$$

y si pasamos la primera fila al lugar n -ésimo, el resultado toma la forma:

$$\begin{vmatrix} P_{k-1}(y) & P_0(y) & P_1(y) & \dots & P_{k-2}(y) \\ P_{k-2}(y) & yP_{k-1}(y) & P_0(y) & P_1(y) & \dots & P_{k-3}(y) \\ \vdots & yP_{k-2}(y) & \ddots & \ddots & \ddots & \ddots \\ P_1(y) & \vdots & \ddots & \ddots & \ddots & \ddots \\ P_0(y) & yP_1(y) & yP_2(y) & \dots & \dots & yP_{k-1}(y) \end{vmatrix} = 0.$$

A modo de ejercicio, se puede encontrar este resultado por el método de Cayley.

§ VI.4 DISCRIMINANTE DE UN POLINOMIO

DEFINICIÓN VI.4.1

§ Sea $f(x)$ un polinomio. El **discriminante** de $f(x)$ es la resultante $R_{f,f'}$ de f y de su derivada respecto de x . Se le designa por $\Delta(f)$.

Si f es un polinomio «general»: $f(x) = u_0 x^m + u_1 x^{m-1} + \dots + u_m$, en donde los u_i son variables, $\Delta(f)$ es un polinomio respecto de las variables u_i . Se puede demostrar que $\Delta(f)$ es entonces un polinomio irreducible a menos del factor u_0 .

(Nota. Este hecho no se sigue del teorema VI.1.3, pues existen relaciones entre los coeficientes de f y de f' .)

Ejemplos

1) Si $f(x) = ax^2 + bx + c$, $f'(x) = 2ax + b$:

$$\Delta(f) = R_{f,f'} = \begin{vmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{vmatrix} = a(4ac - b^2).$$

2) Si $f(x) = ax^3 + bx^2 + cx + d$, $f'(x) = 3ax^2 + 2bx + c$.

El método de Cayley nos da:

$$\Delta(f) = \begin{vmatrix} ab & 2ac & 3ad \\ 2ac & 3ad+bc & 2bd \\ 3a & 2b & c \end{vmatrix},$$

o sea $\Delta(f) = a(18abcd + b^2c^2 - 4b^3d - 4ac^3 - 27a^2d^2)$.

TEOREMA VI.4.1

Sea $f(x) = a_0 x^m + a_1 x^{m-1} + \dots + a_m$ un polinomio con coeficientes en K , y sean $\alpha_1, \alpha_2, \dots, \alpha_m$ sus raíces. Se tiene entonces:

$$(1) \quad \Delta(f) = (-1)^{m(m-1)/2} \left[\prod_{1 \leq i < j \leq m} (\alpha_i - \alpha_j)^2 \right] \times a_0^{2m-1}.$$

Demostración. Se tiene: $f(x) = a_0 \prod_{1 \leq i \leq m} (x - \alpha_i)$, de donde:

$$f'(x) = a_0 \sum_{i=1}^m (x - \alpha_1) \dots (x - \alpha_{i-1}) (x - \alpha_{i+1}) \dots (x - \alpha_m) = \sum_{i=1}^m \frac{f(x)}{x - \alpha_i}.$$

Sabemos (teorema VI.1.4) que

$$(2) \quad R_{f,f'} = a_0^{m-1} \prod_{i=1}^m f'(\alpha_i),$$

de donde:

$$\begin{aligned} \Delta(f) &= a_0^{2m-1} \prod_{i=1}^m ((\alpha_i - \alpha_1) \dots (\alpha_i - \alpha_{i-1}) (\alpha_i - \alpha_{i+1}) \dots (\alpha_i - \alpha_m)) \\ &= a_0^{2m-1} \prod_{i=1}^m (-1)^{i-1} (\alpha_1 - \alpha_i) \dots (\alpha_{i-1} - \alpha_i) (\alpha_i - \alpha_{i+1}) \dots (\alpha_i - \alpha_m) \\ &= (-1)^{m(m-1)/2} a_0^{2m-1} \prod_{i < j} (\alpha_i - \alpha_j)^2. \text{ c.q.d.} \end{aligned}$$

Las fórmulas (1), (2) permiten calcular $\Delta(f)$ como función simétrica de las raíces del polinomio $f(X)$.

Ejemplo

Calculemos $\Delta(f)$ para $f = X^m + pX + q$. Si $\alpha_1, \dots, \alpha_m$ designan las raíces de f , se tiene:

$$\Delta(f) = \prod_{i=1}^m f'(\alpha_i) = \prod_{i=1}^m (m\alpha_i^{m-1} + p).$$

El cálculo se puede abreviar si se tiene en cuenta que $\alpha_i^m = -p\alpha_i - q$, de donde:

$$\alpha_i^{m-1} = -p - \frac{q}{\alpha_i}.$$

Se obtiene:

$$\begin{aligned}\Delta(f) &= \prod_{i=1}^m \left[-mp - \frac{mq}{\alpha_i} + p \right] = (-1)^m \prod_{i=1}^m \left[(m-1)p + \frac{mq}{\alpha_i} \right] \\ \Delta(f) &= (-1)^m \prod_{i=1}^m [(m-1)p + mq\beta_i], \quad \text{haciendo: } \beta_i = \frac{1}{\alpha_i}. \\ &= (-1)^m m^m q^m \prod_{i=1}^m \left[\frac{(m-1)p}{mq} + \beta_i \right].\end{aligned}$$

Luego la ecuación cuyas raíces son las β_i es:

$$qY^m + pY^{m-1} + 1 = 0.$$

Así pues, la ecuación cuyas raíces son las $\beta_i + \frac{(m-1)p}{mq} = \gamma_i$ es:

$$q \left[Z - \frac{(m-1)p}{mq} \right]^m + p \left[Z - \frac{(m-1)p}{mq} \right]^{m-1} + 1 = 0.$$

El producto de las raíces de esta última ecuación se obtiene calculando el término constante:

$$\prod_{i=1}^m \gamma_i = \frac{(-1)^m}{q} \left\{ (-1)^m q \frac{(m-1)^m p^m}{m^m q^m} + (-1)^{m-1} p \frac{(m-1)^{m-1} p^{m-1}}{m^{m-1} q^{m-1}} + 1 \right\}.$$

Se deduce:

$$\boxed{\Delta(f) = (-1)^{m-1} p^m (m-1)^{m-1} + m^m q^{m-1}}.$$

El método de reducción de grado conduce de una manera más rápida al mismo resultado.

Se debe eliminar x entre

$$(2) \quad x^m + px + q = 0 \quad \text{y} \quad (3) \quad mx^{m-1} + p = 0.$$

Multipliquemos (3) por x y (2) por m , restemos y simplifiquemos:

$$-px = +mx^m, \quad m(x^m + px + q) = 0, \quad \text{de donde } p(m-1)x + mq = 0.$$

Llevando a (3) $x = -\frac{mq}{p(m-1)}$, se obtiene la condición necesaria:

$$(-1)^{m-1} \frac{m^m q^{m-1}}{p^{m-1}(m-1)^{m-1}} + p = 0.$$

Puesto que $\Delta(f)$ es de peso $2m-1$ respecto de los coeficientes p, q, m , $\Delta(f)$ es proporcional a: $(-1)^{m-1} m^m q^{m-1} + p^m(m-1)^{m-1}$. Vemos que el coeficiente de proporcionalidad es -1 . c.q.d.

Según la teoría de las raíces múltiples (cf. Cap. IV), *la nulidad de $\Delta(f)$ es necesaria y suficiente para que f posea al menos una raíz múltiple*. Esta propiedad la profundizaremos más adelante y nos conducirá, de forma notable, al teorema que caracteriza las correspondencias algebraicas y biyectivas salvo para un número finito de puntos.

Discriminante y sumas de Newton

Consideremos la ecuación $f(X) = 0$, en donde $f(X) \equiv X^m + a_1 X^{m-1} + \dots + a_m$, con coeficientes en K , y sea $f(x) = \prod_{k=1}^m (X - \alpha_k)$. Designemos por S_k ($k \geq 0$) las sumas de las potencias k -ésimas de estas raíces:

$$S_k = \sum_{i=1}^m \alpha_i^k.$$

Vamos a dar una expresión de $\Delta(f)$ con la ayuda de los S_k . Para ello, consideremos el determinante de Vandermonde:

$$D = \begin{vmatrix} 1 & \alpha_1 & \dots & \alpha_1^{m-1} \\ 1 & \alpha_2 & \dots & \alpha_2^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_m & \dots & \alpha_m^{m-1} \end{vmatrix}.$$

En virtud de la fórmula del desarrollo es: $D = \prod_{1 \leq i < j \leq m} (\alpha_j - \alpha_i)$, se ve que:

$$D^2 = (-1)^{m(m-1)/2} \Delta(f).$$

Sea V la matriz:

$$\begin{bmatrix} 1 & \dots & 1 \\ \alpha_1 & & \alpha_m \\ \vdots & & \vdots \\ \alpha_1^{m-1} & \dots & \alpha_m^{m-1} \end{bmatrix},$$

se tiene:

$$D^2 = (\det V)^2 = \det(V) \times \det({}^tV) = \det(V \times {}^tV).$$

Efectuando el producto $V \times {}^tV$, resulta:

$$(2) \quad \Delta(f) = (-1)^{m(m-1)/2} \begin{vmatrix} S_0 & S_1 & \dots & S_{m-1} \\ S_1 & S_2 & \dots & S_m \\ \vdots & & & \vdots \\ S_{m-1} & S_m & \dots & S_{2m-2} \end{vmatrix}.$$

Ejemplo

Calculemos $\Delta(f)$ para $f(x) = x^5 + px^2 + q$.

Aquí se tiene:

$$S_0=5, \quad S_1=0, \quad S_2=0, \quad S_3=-3p, \quad S_4=0, \quad S_5=-pS_2-5q=-5q, \\ S_6=-pS_3-qS_1=3p^2, \quad S_7=-pS_4-qS_2=0, \quad S_8=-pS_5-qS_3=8pq;$$

de donde

$$\Delta(f) = \begin{vmatrix} 5 & 0 & 0 & -3p & 0 \\ 0 & 0 & -3p & 0 & -5q \\ 0 & -3p & 0 & -5q & 3p^2 \\ -3p & 0 & -5q & 3p^2 & 0 \\ 0 & -5q & 3p^2 & 0 & 8pq \end{vmatrix} = 5^5 q^4 + 108 p^5 q.$$

Téngase en cuenta que, incluso en este caso, el método de reducción del grado es más rápido.

* § VI.5 EXPRESIÓN DE LAS RAÍCES COMUNES A DOS ECUACIONES, CUANDO SU RESULTANTE ES NULA

Sean

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_0, \quad a_m \neq 0 \\ g(x) = b_n x^n + b_{n-1} x^{n-1} + \dots + b_0, \quad b_n \neq 0$$

dos polinomios con coeficientes en K . Designemos por $\mathbf{Z}[a, b]$ al anillo $\mathbf{Z}[a_0, \dots, a_m; b_0, \dots, b_n]$, y por $\mathbf{Q}(a, b)$ al cuerpo $\mathbf{Q}(a_0, \dots, a_m; b_0, \dots, b_n)$. $\mathbf{Q}(a, b)$ es el cuerpo de fracciones de $\mathbf{Z}[a, b]$.

La teoría del mcd (algoritmo de Euclides) nos demuestra que

$$D(x) = \text{mcd}(f(x), g(x))$$

es un polinomio con coeficientes en $\mathbf{Q}(a, b)$.

En particular, si $R(f, g) = 0$, y si $\text{mcd}(f, g)$ es exactamente de grado 1, la raíz común a f, g es un elemento de $\mathbf{Q}(a, b)$, lo que se enuncia diciendo que, en este caso, la raíz común a f y g se expresa racionalmente en función de los coeficientes de f y g .

Por el contrario, si el mcd de f y de g es de grado > 1 , esta propiedad puede ser falsa: por ejemplo, las raíces comunes a $x^3 - x^2 - 2x + 2$ y a $x^2 - 2$ son $\pm \sqrt{2}$. En consecuencia nos encontramos con los dos problemas siguientes:

- (I) Reconocer, según los coeficientes de f y g , el grado exacto del $\text{mcd}(f, g)$.
- (II) Conociendo el grado del $\text{mcd}(f, g)$, expresar este polinomio con la ayuda de una fórmula «global» que haga intervenir los coeficientes de f y g .

La solución de (II) nos proporcionará en particular la expresión de la raíz común a f y g cuando el grado del $\text{mcd}(f, g)$ es de grado 1.

La solución práctica de los problemas (I) y (II) consiste en formar el mcd de f y g por los métodos ya conocidos (por ejemplo, la reducción del grado — que se reduce al algoritmo de Euclides). Hemos visto ejemplos y no insistiremos más acerca de este método.

Solución teórica del problema (I)

Los polinomios considerados a continuación son elementos de $K[X]$.

Suponiendo conocido un divisor común D , de grado p , de f, g , por ejemplo,

$$D(X) = (X - \alpha_1)(X - \alpha_2) \dots (X - \alpha_p) = X^p - \sigma_1 X^{p-1} + \sigma_2 X^{p-2} + \dots + (-1)^p \sigma_p,$$

vamos a formar la resultante de $f_1 = \frac{f}{D}$ y $g_1 = \frac{g}{D}$, en función de los coeficientes de f y g , con:

$$f(X) = \sum_{k=0}^m a_k X^{m-k}, \quad g(X) = \sum_{k=0}^n b_k X^{n-k}.$$

Para ello introducimos una matriz A_0 , deducida por permutación de las columnas de la matriz de Sylvester de f y g . Con más precisión, A_0 es la matriz de los polinomios

$$X^{n-1}f(X), X^{n-2}f(X), \dots, Xf(X), f(X), g(X), Xg(X), \dots, X^{m-1}g(X)$$

en la base $(X^{m+n-1}, \dots, X, 1)$,

$$A_0 = \begin{bmatrix} a_0 & 0 & \dots & 0 & b_0 \\ a_1 & a_0 & \dots & 0 & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_m & a_{m-1} & \dots & a_0 & b_n \\ 0 & a_m & \dots & a_{m-1} & b_{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

$\leftarrow A_1$
 $\leftarrow A_2$

Designamos por A_k a la matriz formada por los elementos de A cuyos índices de filas y columnas están estrictamente comprendidos entre $m+n-k$ y k , para $k \leq \min(m, n)$. El resultado esencial es el siguiente:

TEOREMA VI.5.1

$$\left\| \begin{array}{l} \text{Con las notaciones que preceden, la resultante de } f_1(X) \text{ y } g_1(X) \text{ es } (-1)^{\frac{m(m-1)}{2}} \\ \det(A_p). \end{array} \right.$$

Utilizaremos dos resultados preliminares:

$$\left\| \begin{array}{l} \text{VI.5.2 Sean } f = Df_1 \text{ y } g = Dg_1 \text{ dos polinomios divisibles por } D. \text{ Para que } f_1 \text{ y } g_1 \\ \text{tengan raíces comunes, es necesario y suficiente que existan dos polino-} \\ \text{mios } U \text{ y } V \text{ tales que: } U \neq 0, V \neq 0 \text{ y:} \\ Uf + Vg = 0, \quad \text{gr}(U) \leq n - p - 1, \quad \text{gr}(V) \leq m - p - 1. \end{array} \right.$$

Demostración. Es un calco de la de VI.1.1. c.q.d.

El enunciado VI.5.2 significa que los polinomios

$$X^{n-p-1}f, \dots, Xf, f, g, Xg, \dots, X^{m-p-1}g$$

deben estar ligados, y por tanto que la matriz B_p formada con las columnas de A de índice estrictamente comprendido entre p y $m+n-p$, es de rango $\leq m+n-2p-1$.

VI.5.3 Sea $D = (X - \alpha_1) \dots (X - \alpha_p)$ un polinomio fijo. El espacio vectorial de los polinomios con coeficientes en K y grado $\leq m$, que son múltiplos de D , es de dimensión $m - p + 1$. Con mas precisión, estos polinomios son de la forma:

$$f(X) = a_0 X^m + a_1 X^{m-1} + \dots + a_{m-p} X^p + a_{m-p+1} X^{p-1} + \dots + a_m,$$

en donde a_0, a_1, \dots, a_{m-p} son arbitrarios, y en donde a_{m-p+1}, \dots, a_m se hallan expresados en función de a_0, a_1, \dots, a_{m-p} por medio de formas lineales **fijas**.

Demostración. Basta con escribir: $D = X^p - \sigma_1 X^{p-1} + \dots + (-1)^p \sigma_p$, $f(X) = (X^p - \sigma_1 X^{p-1} + \dots + (-1)^p \sigma_p) (\lambda_0 X^{m-p} + \lambda_1 X^{m-p-1} + \dots + \lambda_{m-p})$, de donde

$$(1) \quad \begin{cases} a_0 &= \lambda_0 \\ a_1 &= -\sigma_1 \lambda_0 + \lambda_1 \\ a_2 &= \sigma_2 \lambda_0 - \sigma_1 \lambda_1 + \lambda_2 \\ \vdots & \\ a_{m-p} &= (-1)^p \sigma_p \lambda_0 + (-1)^{p-1} \sigma_{p-1} \lambda_1 + \dots + \lambda_{m-p} \end{cases}$$

$\lambda_0, \lambda_1, \dots, \lambda_{m-p}$ son arbitrarios, luego a_0, a_1, \dots, a_{m-p} son arbitrarios, pues las fórmulas (1) son invertibles (recordemos que los σ_i son *fijos*). Las otras fórmulas obtenidas por la identificación de f y $D(\lambda_0 X^{m-p} + \dots + \lambda_{m-p})$, que son:

$$(2) \quad a_{m-p+1} = (-1)^p \sigma_p \lambda_1 + (-1)^{p-1} \sigma_{p-1} \lambda_2 + \dots, \text{ etc.,}$$

dan los a_{m-p+i} ($i \geq 1$) en función de los (λ_i) , y, por lo tanto, también en función de los a_k ($k \leq m-p$). c.q.d.

Demostración del teorema VI.5.1. Fijamos D , y hacemos variar arbitrariamente f_1 y g_1 .

La resultante $R_1(f_1, g_1)$ es un polinomio homogéneo de grado $m + n - 2p$ respecto de los coeficientes de f_1 y g_1 , a la que llamaremos simplemente R_1 . Luego, según las relaciones (1) (que permiten obtener los coeficientes de $f_1 = \frac{f}{D}$ y $g_1 = \frac{g}{D}$ en función de los de f y g), existe un polinomio homogéneo $T_1 \in K[X_0, \dots, X_{m-p}; Y_0, \dots, Y_{n-p}]$, de grado $m + n - 2p$, tal que

$$R_1(f_1, g_1) = T_1(a_0, a_1, \dots, a_{m-p}, b_0, b_1, \dots, b_{n-p}).$$

Puesto que T_1 se deduce de R_1 por un cambio de variables lineal e invertible, la irreducibilidad de R_1 implica la de T_1 .

En la matriz A_p , sustituimos $a_0, \dots, a_{m-p}, b_0, \dots, b_{n-p}$ por las variables $X_0, \dots, X_{m-p}, Y_0, \dots, Y_{n-p}$, y se obtiene una matriz M_p , cuyo determinante Δ_p es un polinomio nulo u homogéneo de grado $m + n - 2p$ en las X_i y las Y_j . Se tiene: $\Delta_p \neq 0$, pues Δ_p contiene el monomio $\varepsilon X_0^{m-p} Y_{n-p}^{n-p}$ (en donde $\varepsilon = \pm 1$).

Sean $\varphi_1, \dots, \varphi_p$ las formas lineales introducidas en VI.5.3.

Según VI.5.3, poniendo:

$$\begin{aligned} F_p(X_0, \dots, X_{m-p}, Y_0, \dots, Y_{n-p}) \\ = \Delta_p(X_0, \dots, X_{m-p}; \Phi_1, \dots, \Phi_p; Y_0, \dots, Y_{n-p}, \Psi_1, \dots, \Psi_p) \end{aligned}$$

(en donde $\Psi_i = \varphi_i(Y_0, \dots, Y_{n-p})$ y $\Phi_i = \varphi_i(X_0, \dots, X_{m-p})$ para $1 \leq i \leq p$) se tiene:

$$\det(A_p) = F_p(a_0, \dots, a_{m-p}; b_0, \dots, b_{n-p}).$$

El polinomio F_p es no nulo, pues, dando a los (a_i) y a los (b_j) ($0 \leq i \leq m-p$, $0 \leq j \leq n-p$) valores tales que

$$f(X) = a_0 X^{m-p} D(X), \quad g(X) = b_{n-p} D(X),$$

el valor tomado por $\det(A_p)$ es $(-1)^{(m-p)(m-p-1)/2} a_0^{n-p} b_{n-p}^{m-p}$. Luego $\text{gr}(F_p) = m + n - 2p$.

Finalmente, sean $a_0, \dots, a_{m-p}; b_0, \dots, b_{n-p}$ escalares elegidos en K de forma que $T_1(a_0, \dots, a_{m-p}, b_0, \dots, b_{n-p}) = 0$.

Entonces, según la nota que sigue a VI.5.2, F_p se anula para estos mismos valores. Luego F_p es proporcional a T_1 . Dando a los (a_i) y a los (b_j) valores tales que $f(X) = X^{m-p}$ y $g(X) = b_{n-p}$, vemos que el factor de proporcionalidad es $(-1)^{m(m-1)/2}$.

El problema (I) está, pues, resuelto: *para que el mcd (f, g) sea exactamente de grado p (en donde $1 \leq p \leq \inf(m, n)$), es necesario y suficiente que*

$$\det(A_0) = \det(A_1) = \dots = \det(A_{p-1}) = 0,$$

y que $\det(A_p) \neq 0$.

Ejemplo

Busquemos las condiciones para que $f(T) = T^3 + xT^2 + yT + z$ posea una raíz triple. Es suficiente expresar que el mcd de $f(T)$ y $f'(T)$ es de grado 2, puesto que f no puede admitir dos raíces dobles distintas.

La matriz A_0 es:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ x & 1 & 0 & 3 & 2x \\ y & x & 3 & 2x & y \\ z & y & 2x & y & 0 \\ 0 & z & y & 0 & 0 \end{bmatrix}.$$

Sabemos que el $\det(A_0)$ es el discriminante $\Delta(f)$ de f :

$$\Delta(f) = 18xyz + x^2y^2 - 4x^3z - 4y^3 - 27z^2.$$

Además, $A_1 = \begin{bmatrix} 1 & 0 & 3 \\ x & 3 & 2x \\ y & 2x & y \end{bmatrix}$, y la condición $\det(A_1) = 0$ expresa, como debe

ser, que $f'(T)$ posee una raíz doble: $x^2 - 3y = 0$.

Las condiciones buscadas equivalen al sistema

$$\begin{aligned} (2) \quad & \begin{cases} 6xyz - y^3 - 27z^2 = 0 \\ (3) \quad \begin{cases} x^2 - 3y = 0. \end{cases} \end{cases} \end{aligned}$$

Estas condiciones equivalen también a la existencia de un t tal que $f(T) \equiv (T+t)^3$, es decir:

$$(4) \quad x = 3t, \quad y = 3t^2, \quad z = t^3;$$

éstas son las ecuaciones paramétricas de una *cúbica alabeada* de K^3 , que designaremos por (I') .

La eliminación directa de t entre las relaciones (4) nos habría conducido a las relaciones

$$(5) \quad x^2 - 3y = 0, \quad (6) \quad 27z = x^3.$$

Sin embargo, las relaciones (2) y (3) presentan, desde cierto punto de vista, un mayor interés que (5) y (6).

En efecto, la cúbica alabeada (I') está contenida en la intersección del cilindro parabólico (5) y de la superficie cúbica (6) en K^3 . Pero en el espacio

proyectivo $\mathcal{P}_3(K)$ estas dos superficies poseen también una recta común en el infinito (a lo largo de la cual son tangentes).

(I) es también la intersección de la superficie cúbica (2) y del cilindro parabólico (3). Pero esta vez, (I) es *intersección completa* de estas superficies. Se comprueba fácilmente (¡hacer el dibujo!) que las dos superficies son *tangentes a lo largo de I* .

Estudio del problema (II)

Expondremos el método mediante un ejemplo. Supongamos

$$f(X) = a_0 X^5 + a_1 X^4 + a_2 X^3 + a_3 X^2 + a_4 X + a_5$$

$$g(X) = b_0 X^3 + b_1 X^2 + b_2 X + b_3,$$

y que $D = \text{mcd}(f, g)$ sea de grado 1. Consideremos el determinante:

$$\Delta = \begin{vmatrix} a_0 & 0 & 0 & 0 & 0 & b_0 \\ a_1 & a_0 & 0 & 0 & b_0 & b_1 \\ a_2 & a_1 & 0 & b_0 & b_1 & b_2 \\ a_3 & a_2 & b_0 & b_1 & b_2 & b_3 \\ a_4 & a_3 & b_1 & b_2 & b_3 & 0 \\ xf(x) & f(x) & g(x) & xg(x) & x^2 g(x) & x^3 g(x) \end{vmatrix}$$

formado de forma evidente con la matriz A_1 de f y g (notaciones del teorema VI.5.1). Sean L la última fila de Δ , y L_1, L_2, L_3, L_4, L_5 las 5 primeras filas. Se tiene:

$$(7) \quad x^6 L_1 + x^5 L_2 + x^4 L_3 + x^3 L_4 + x^2 L_5 = L - N,$$

con

$$N = [a_5 x, a_4 x + a_5, b_2 x + b_3, b_3 x, 0, 0].$$

Si calculamos Δ teniendo en cuenta (7) vemos que Δ es un polinomio de grado ≤ 1 en x . Por otra parte, si desarrollamos Δ directamente según la última fila, obtenemos una relación de la forma $\Delta = Uf + Vg$, en donde U y V son polinomios tales que $\text{gr}(U) \leq 1$ y $\text{gr}(V) \leq 3$. Según la teoría de los polinomios, se sigue que Δ es el mcd de f y g , a menos que Δ sea nulo (ya que el grado de Δ es ≤ 1).

Pero si $\Delta = 0$, se tiene necesariamente $U = V = 0$, si no se aplica VI.5.2 y se deduce que el mcd de f y g es de grado > 1 , contrariamente a la hipótesis.

Luego los coeficientes de U y V son los menores de los elementos de la última fila de la matriz A_1 . Según VI.5.1 estos menores no son todos nulos, y no es, pues, posible tener $U = V = 0$, luego Δ es no nulo y, por lo tanto, es el mcd de f y g .

De forma general, se tiene:

TEOREMA VI.5.4

Sean $f(X) = \sum_{k=0}^m a_k X^k$ y $g(X) = \sum_{k=0}^n b_k X^k$ dos polinomios cuyo mcd es de grado p . El mcd de f y g es el determinante de la matriz obtenida reemplazando, en la matriz A_p del teorema VI.5.1, la última fila por los polinomios:

$$x^{n-p-1}f(x), \quad x^{n-p-2}f(x), \dots, xf(x), \quad f(x), \quad g(x), \\ xg(x), \dots, x^{m-p-1}g(x).$$

Ejemplo

Cuando $f(x) = \sum_{k=0}^5 a_k x^k$ y $g(x) = \sum_{k=0}^3 b_k x^k$ tienen exactamente una raíz común, su mcd es:

$$D = \begin{vmatrix} a_0 & 0 & 0 & 0 & 0 & b_0 \\ a_1 & a_0 & 0 & 0 & b_0 & b_1 \\ a_2 & a_1 & 0 & b_0 & b_1 & b_2 \\ a_3 & a_2 & b_0 & b_1 & b_2 & b_3 \\ a_4 & a_3 & b_1 & b_2 & b_3 & 0 \\ a_5 x & a_4 x + a_5 & b_2 x + b_3 & xb_3 & 0 & 0 \end{vmatrix}$$

su raíz común es, pues, $\alpha = -\frac{D_1}{\Delta_1}$, con

$$D_1 = \begin{vmatrix} a_0 & 0 & 0 & 0 & 0 & b_0 \\ a_1 & a_0 & 0 & 0 & b_0 & b_1 \\ a_2 & a_1 & 0 & b_0 & b_1 & b_2 \\ a_3 & a_2 & b_0 & b_1 & b_2 & b_3 \\ a_4 & a_3 & b_1 & b_2 & b_3 & 0 \\ 0 & a_5 & b_3 & 0 & 0 & 0 \end{vmatrix}, \quad \Delta_1 = \begin{vmatrix} a_0 & 0 & 0 & 0 & 0 & b_0 \\ a_1 & a_0 & 0 & 0 & b_0 & b_1 \\ a_2 & a_1 & 0 & b_0 & b_1 & b_2 \\ a_3 & a_2 & b_0 & b_1 & b_2 & b_3 \\ a_4 & a_3 & b_1 & b_2 & b_3 & 0 \\ a_5 & a_4 & b_2 & b_3 & 0 & 0 \end{vmatrix}.$$

Capítulo VII

Fracciones racionales

§ VII.1 EL CUERPO $K(X)$

● En lo que sigue, K designa un cuerpo conmutativo cualquiera. En el capítulo III, § 6 hemos visto como, a partir de un anillo íntegro conmutativo cualquiera, se puede formar su *cuerpo de fracciones*. Esta construcción se aplica al anillo $K[X]$ de los polinomios en una variable, y recordemos (III.6.5) que la relación $\frac{P_1}{Q_1} = \frac{P_2}{Q_2}$ significa: $P_1 Q_2 = P_2 Q_1$.

DEFINICIÓN VII.1.1

Supongamos que K designa un cuerpo conmutativo. Se llama **cuerpo de las fracciones racionales con coeficientes en K** , al cuerpo de las fracciones del anillo de los polinomios $K[X]$. A los elementos de este cuerpo se les llama **fracciones racionales**.

Notación: El cuerpo ahora definido, se designa por $K(X)$.
Recordemos que $K(X)$ se obtiene como conjunto cociente de

$$\mathcal{K} = K[X] \times K[X]^*$$

por la relación de equivalencia \mathcal{R} : $(A, B) \mathcal{R} (A', B')$ si, y sólo si, $AB' = BA'$, el conjunto \mathcal{K} se halla dotado de las dos leyes (compatibles con \mathcal{R})

$$(A, B) + (A', B') = (AB' + BA', BB') \quad (\text{adición})$$

$$(A, B) \cdot (A', B') = (AA', BB') \quad (\text{multiplicación}),$$

El conjunto $K(X) = \mathcal{K}/\mathcal{R}$ está dotado de la estructura de anillo definida por las leyes cociente, y es un *cuerpo*.

La clase del elemento (A, B) se designa por A/B o $\frac{A}{B}$.

La aplicación $P \mapsto \frac{P}{1}$ es un homomorfismo inyectivo de anillos de $K[X]$ en $K(X)$, por medio del cual se identifican los elementos de $K[X]$ con los de un subanillo de $K(X)$. El elemento nulo es $\frac{0}{1}$ (designado por 0).

El elemento unidad de $K(X)$ es 1; el inverso de $\frac{A}{B}$ ($A \neq 0, B \neq 0$) es $\frac{B}{A}$.

Forma irreducible

El hecho de que $K[X]$ sea un anillo *principal* nos permitirá distinguir, entre los representantes posibles de la fracción $F = \frac{A}{B}$, una forma notable. En forma precisa, si designamos por Δ el mcd de A y B , se tiene: $A = \Delta P$, $B = \Delta Q$, con P y Q primos entre sí, de donde $F = \frac{P}{Q}$. Sean P_1 y Q_1 otro par de polinomios primos entre sí tales que $F = \frac{P_1}{Q_1}$. De la relación $PQ_1 = P_1Q$, y del teorema de Gauss (Cap. III, § 2) deducimos que P divide a P_1 y que P_1 divide a P . En otras palabras, P y P_1 están *asociados* e, igualmente, Q y Q_1 están asociados. Existe un par y sólo uno (P, Q) de polinomios primos entre sí, tales que $F = \frac{P}{Q}$ (salvo un factor constante no nulo común a P y Q). Enunciaremos:

VII.1.1 Para cada fracción racional $F \in K(X)$, existe un representante $\frac{P}{Q}$ de F , tal que los polinomios P y Q sean primos entre sí. Este representante es único salvo factores constantes no nulos, y se le llama **forma irreducible** de F .

Sea $F = \frac{P}{Q}$ una fracción racional expresada en forma irreducible, y sea $F = \frac{A}{B}$ otra representación de F . De la relación $AQ = BP$, y del teorema de Gauss, deducimos la existencia de un polinomio D tal que $A = DP$ y $B = DQ$. Este resultado

se puede expresar diciendo que toda representación $\frac{A}{B}$ de la fracción F es un «múltiplo» de su forma irreducible.

Funciones racionales

Sea L un subcuerpo conmutativo de K , que puede ser el mismo K , y consideremos una fracción $F \in K(X)$, expresada en forma irreducible: $F = \frac{P}{Q}$.

DEFINICIÓN VII.1.2

El conjunto de las raíces en L del denominador Q de una fracción racional F , expresada en forma irreducible, se llama **conjunto de los polos de F en L** , y es un conjunto **finito**. Al complementario, en L , del conjunto de los polos de F , se le llama **conjunto de definición de F en L** .

Designamos por $D_{F,L}$ al conjunto de definición de F en L . Para todo $x \in D_{F,L}$, se tiene $Q(x) \neq 0$, luego el elemento $\frac{P(x)}{Q(x)} \in L$ está bien definido.

DEFINICIÓN VII.1.3

Con las notaciones anteriores, a la aplicación $\tilde{F} : D_{F,L} \rightarrow L$, tal que, para todo $x \in D_{F,L}$, $\tilde{F}(x) = \frac{P(x)}{Q(x)}$, se le llama **función racional asociada a F en L** .

En ciertos casos, el conjunto de definición puede ser vacío. Por ejemplo, si K es un cuerpo finito, de elementos (a_1, a_2, \dots, a_q) , y ponemos:

$$Q(X) = (X - a_1)(X - a_2) \dots (X - a_q).$$

El conjunto de definición en K de la fracción $F = \frac{1}{Q}$ es vacío, pues todo elemento de K es polo de F .

Designemos por $F = \frac{P}{Q}$ y $G = \frac{R}{S}$ dos fracciones expresadas en forma irreducible. Para todo elemento x común a los conjuntos de definición de F y G , se tiene evidentemente:

$$\begin{aligned}\tilde{F}(x) + \tilde{G}(x) &= (\widetilde{F + G})(x), \\ \tilde{F}(x) \cdot \tilde{G}(x) &= (\widetilde{FG})(x).\end{aligned}$$

Finalmente, en la mayor parte de los casos usuales, una fracción racional está determinada por la aplicación racional asociada:

TEOREMA VII.1.2

Sean K un cuerpo conmutativo **infinito** y $F_1 \in K(X)$; $F_2 \in K(X)$. La relación « $\tilde{F}_1(x) = \tilde{F}_2(x)$ para todo x de la intersección de los conjuntos de definición de F_1 y F_2 en K » implica « $F_1 = F_2$ ».

Demostración. Escribamos F_1 y F_2 en forma irreducible:

$$F_1 = \frac{P_1}{Q_1}, \quad F_2 = \frac{P_2}{Q_2}$$

La intersección D de los conjuntos de definición de F_1 y F_2 es una parte infinita de K , puesto que el conjunto de los polos de F_1 y F_2 es finito. Para todo $x \in D$, se tiene:

$$P_1(x) Q_2(x) = P_2(x) Q_1(x).$$

Los polinomios $P_1 Q_2$ y $P_2 Q_1$, iguales para una infinidad de valores de x , son formalmente iguales. c.q.d.

Aplicación

Si K es de característica nula (y contiene, por lo tanto, a \mathbf{Q}),

$$(\tilde{F}_1(x) = \tilde{F}_2(x) \text{ para todo } x \in \mathbf{Q}) \Rightarrow (F_1 = F_2).$$

Multiplicidad de un polo

Consideremos de nuevo la fracción $F = \frac{P}{Q} \in K(X)$, expresada en forma irreducible, y el subcuerpo conmutativo L de K . Si a designa un polo de F en L , por definición, la *multiplicidad del polo* a es la multiplicidad de a como raíz de Q . Cuando esta multiplicidad es 1, se dice que a es un *polo simple* de F , cuando es 2, 3, ..., n , se dice que a es un polo *doble*, *triple*, ..., *n-plo*.

El cuerpo $K(X_1, \dots, X_n)$

Por definición, el cuerpo de las fracciones del anillo $K(X_1, \dots, X_n)$ (K , cuerpo conmutativo) es el cuerpo de las fracciones racionales con n variables sobre el

cuerpo K . En este contexto tiene sentido también hablar de *forma irreducible* de una fracción: en efecto, para obtenerla en el caso $n = 1$, hemos utilizado únicamente el *teorema de Gauss*, y no el hecho de que $K[X]$ sea principal. Ahora bien, el teorema de Gauss es válido también en el anillo $K[X_1, X_2, \dots, X_n]$ (cf. Cap. XIV).

Es posible, pues, en el caso de las fracciones con n variables, definir los conjuntos de polos y los conjuntos de definición, exactamente como se ha hecho anteriormente. Pero ahora, el conjunto de polos es menos simple que en el caso $n = 1$ (es una «hipersuperficie» algebraica de K^n).

El teorema análogo al teorema VI.1.1 es válido también en $K(X_1, \dots, X_n)$, pero establecerlo es más delicado. Se basa en el teorema siguiente (Teorema XIV.4.5, p. 520):

«Si dos polinomios con n variables toman los mismos valores sobre el complementario de una hipersuperficie algebraica de K^n , son formalmente iguales».

Observemos, finalmente, que los polos de una fracción $F = \frac{P(X_1, \dots, X_n)}{Q(X_1, \dots, X_n)}$ pueden ser de naturaleza muy diversa: un polo $(\alpha_1, \dots, \alpha_n)$ tal que $P(\alpha_1, \dots, \alpha_n) \neq 0$ y $Q(\alpha_1, \dots, \alpha_n) = 0$ es evidentemente mucho menos «complicado» que un polo $(\alpha_1, \dots, \alpha_n)$ tal que $P(\alpha_1, \dots, \alpha_n) = Q(\alpha_1, \dots, \alpha_n) = 0$. Para $n \geq 2$, esta última relación puede realizarse con polinomios P y Q primos entre sí.

§ VII.2 DESCOMPOSICIÓN EN ELEMENTOS SIMPLES SOBRE UN CUERPO CUALQUIERA K

Consideremos la fracción racional $F = \frac{P}{Q}$, expresada en forma irreducible con Q normalizado. Descompongamos el polinomio Q en factores irreducibles:

$$Q = \prod_{i=1}^r Q_i^{\alpha_i} \quad \alpha_i \geq 1,$$

en donde los Q_i son primos entre sí dos a dos, irreducibles y normalizados.

TEOREMA VII.2.1

Con las notaciones anteriores, la fracción F se escribe, de una manera y una sola, en la forma:

$$(1) \quad F = E + \sum_{i=1}^r \mathcal{P}_i,$$

en donde E es un polinomio, y en donde:

$$(2) \quad \mathcal{P}_i = \sum_{j=1}^{\alpha_i} \frac{P_{i,j}}{(Q_i)^j},$$

siendo los polinomios $P_{i,j}$ y los polinomios Q_i tales que, para todo i y todo j , $\text{gr}(P_{i,j}) < \text{gr}(Q_i)$.

Demostración

a) *Existencia.* Ante todo suponemos: $Q = RS$, R y S primos entre sí. Según el teorema de Bezout (IV.3.2) existen polinomios U y V tales que $UR + VS = 1$, por lo que se tiene:

$$F = \frac{P}{Q} = \frac{P(UR + VS)}{RS} = \frac{PU}{S} + \frac{PV}{R}.$$

Aplicando esta fórmula a la fracción $\frac{P}{Q}$, con $R = Q_1^{\alpha_1}$, y observando que $Q_1^{\alpha_1}$ es primo con el producto $Q_2^{\alpha_2} \dots Q_r^{\alpha_r}$, se obtiene en primer lugar la descomposición:

$$\frac{P}{Q} = \frac{PV}{Q_1^{\alpha_1}} + \frac{PU}{Q_2^{\alpha_2} \dots Q_r^{\alpha_r}};$$

paso a paso, aplicando sucesivamente esta misma fórmula con $R = Q_2^{\alpha_2}$, $R = Q_3^{\alpha_3}, \dots$, $R = Q_{r-1}^{\alpha_{r-1}}$, se llega a expresar $F = \frac{P}{Q}$ en la forma

$$(3) \quad F = \sum_{i=1}^r \frac{A_i}{Q_i^{\alpha_i}},$$

en donde los A_i son polinomios.

Admitiendo provisionalmente el resultado posterior VII.2.2, A_i se puede escribir en la forma:

$$A_i = P_{i,\alpha_i} + P_{i,\alpha_i-1} Q_i + \dots + P_{i,1} Q_i^{\alpha_i-1} + E_i Q_i^{\alpha_i},$$

en donde E_i es un polinomio y los $P_{i,j}$ son polinomios de grado $< \text{gr } Q_i$.

Substituyendo, en la fórmula (3), A_i por esta expresión, y tomando $E = \sum E_i$, se obtiene una descomposición de F en la forma enunciada en (1)-(2).

b) *Unicidad*. Quedará establecida de forma evidente si probamos la propiedad siguiente: una relación de la forma

$$(4) \quad E + \sum_{i=1}^r \mathcal{P}_i = 0,$$

en donde E es un polinomio, y en donde $\mathcal{P}_i = \sum_{j=1}^{a_i} \frac{P_{i,j}}{(Q_i)^j}$, con $\text{gr}(P_{i,j}) < \text{gr}(Q_i)$, implica $E = 0$ y $P_{i,j} = 0$ para todo i y todo j .

Hagamos $R_i = \prod_{k \neq i} (Q_k)^{a_k}$. Multipliquemos el primer miembro de (4) por Q_i , y obtendremos:

$$(5) \quad P_{i,\alpha_i} R_i + Q_i S_i = 0,$$

en donde S_i designa un polinomio. Puesto que Q_i es primo y $\text{gr}(P_{i,\alpha_i}) < \text{gr}(Q_i)$, si P_{i,α_i} fuese no nulo P_{i,α_i} y Q_i serían primos entre sí; luego (T. de Gauss) R_i divide a S_i , y (5) implica (si T_i designa un nuevo polinomio)

$$P_{i,\alpha_i} + Q_i T_i = 0.$$

Esto, junto con la desigualdad $\text{gr}(P_{i,\alpha_i}) < \text{gr}(Q_i)$, implica $P_{i,\alpha_i} = 0$.

Sucesivamente, se demuestra de forma análoga que $P_{i,\alpha_{i-2}} = 0$, $P_{i,\alpha_{i-1}} = 0$, etc. Así todos los $P_{i,j}$ son nulos, y se deduce que $E = 0$, c.q.d.

Hemos utilizado, sobre la marcha, el resultado siguiente, que merece ser enunciado aparte (y que podría, además, generalizarse a anillos más generales).

VII.2.2 Sean A y P dos polinomios cualesquiera de $K[X]$, $P \neq 0$. Para todo entero n , existe un sistema, y sólo uno, de polinomios $A_0, A_1, \dots, A_{n-1}, R_n$ tales que:

$$(6) \quad A = A_0 + A_1 P + A_2 P^2 + \dots + A_{n-1} P^{n-1} + R_n P^n,$$

$$\text{y } \text{gr}(A_k) < \text{gr}(P) \quad (0 \leq k \leq n-1).$$

Demostración

a) *Unicidad*. Suponemos que se cumple:

$$A_0 + A_1 P + A_2 P^2 + \dots + A_{n-1} P^{n-1} + R_n P^n = 0.$$

Escribiendo:

$$A_0 = -P(A_1 + A_2 P + \dots + R_n P^{n-1}),$$

y teniendo en cuenta que $\text{gr}(A_0) < \text{gr}(P)$, se ve que $A_0 = 0$. Tras haber simplificado por P la relación anterior, se obtendrá asimismo $A_1 = 0$, etc., $A_{n-1} = 0$ y $R_n = 0$.

b) *Existencia*. Se prueba por recurrencia sobre n . Para $n = 1$, la propiedad se reduce a la división euclídea de A por P . Supongamos obtenida la expresión (6) hasta n . La división euclídea de R_n por P nos da:

$$R_n = R_{n+1}P + A_n, \quad \text{gr}(A_n) < \text{gr}(P),$$

de donde

$$\begin{aligned} A &= A_0 + A_1P + \cdots + A_{n-1}P^{n-1} + (A_n + R_{n+1}P)P^n \\ &= \sum_{k=0}^n A_k P^k + R_{n+1}P^{n+1} \quad \text{c.q.d.} \end{aligned}$$

Nota. Se puede expresar el teorema VI.2.1 en lenguaje de álgebra lineal; designemos por \mathcal{P} al conjunto de los polinomios irreducibles *normalizados* de $K[X]$: dos elementos distintos de \mathcal{P} son pues primos entre sí. El teorema VI.2.1 significa que se ha obtenido una *base del K -espacio vectorial $K(X)$* , a saber el conjunto \mathcal{B} de las fracciones racionales que sigue:

$$(X^n)_{n \geq 0}, \quad \left(\frac{X^k}{P^l} \right) \quad (P \in \mathcal{P}, l \geq 1, 0 \leq k < \text{gr}(P)).$$

(Volviendo a la parte *b*) de la demostración de VII.2.1, se puede, en efecto, demostrar que \mathcal{B} es libre.)

En las fórmulas (1)-(2), al término \mathcal{P}_i se le llama *parte polar de F* relativa al factor Q_{ai} del denominador Q de F , y a E se le llama *parte entera de F* . Multiplicando (1) por Q , se obtiene:

$$P = EQ + R, \quad \text{gr}(R) < \text{gr}(Q).$$

Por consiguiente, E es precisamente el cociente en la división euclídea de P por Q ; y, en los cálculos prácticos, ante todo, interesará determinar E .

● Cuando K es algebraicamente cerrado (en particular si $K = \mathbb{C}$) los factores Q_i son todos ellos de primer grado, o sea $Q_i = x - a_i$, y los polinomios $P_{i,j}$ se reducirán a constantes, y toda fracción racional $F = \frac{P}{Q}$ sobre K admitirá una descomposición única de la forma:

$$(7) \quad \boxed{F = E + \sum_{i=1}^r \sum_{j=1}^{\alpha_i} \frac{A_{i,j}}{(x - a_i)^j}}$$

en donde a_1, a_2, \dots, a_r son las distintas raíces de Q , y $\alpha_1, \alpha_2, \dots, \alpha_r$ sus órdenes de multiplicidad, los $A_{i,j}$ son constantes, y E es un polinomio (cociente de P por Q) de grado igual a $\text{gr}(P) - \text{gr}(Q)$.

Se dice entonces que a_i es un *polo de orden* α_i de F , y, la fracción racional

$$\mathcal{P}_i = \sum_{j=1}^{\alpha_i} \frac{A_{i,j}}{(x - a_i)^j}$$

se, llama *parte polar* de F relativa al polo a_i .

Hagamos, finalmente, una observación que es importante en los cálculos prácticos; si $Q = RS$ (R y S primos entre sí), hemos visto que se verifica:

$$\frac{P}{Q} = F = \frac{A}{R} + \frac{B}{S},$$

en donde A, B son polinomios ⁽¹⁾. Supongamos entonces que el factor $Q_i^{\alpha_i}$ divide a R . Entonces *la parte polar de F relativa al factor $Q_i^{\alpha_i}$, es igual a la parte polar de $\frac{A}{R}$ relativa al factor $Q_i^{\alpha_i}$* , según resulta también de la *unicidad* de la descomposición (1)-(2).

§ VII.3 CÁLCULO DE LAS PARTES POLARES RELATIVAS A LOS FACTORES DE LA FORMA $(X - a)^a$

División según las potencias crecientes

TEOREMA VII.3.1

Sean A, B dos polinomios y se supone $B(0) \neq 0$. Para todo entero n , existe un par de polinomios (Q_n, R_n) , y sólo uno que verifica:

$$(1) \quad A = BQ_n + X^{n+1} R_n, \quad \text{gr}(Q_n) \leq n.$$

A Q_n se le llama *cociente de orden n* , $X^{n+1} R_n$ es el *resto de orden n* .

Demostración

a) *Unicidad*. Si $BQ_n + X^{n+1} R_n = 0$, la condición $B(0) \neq 0$ muestra que X^{n+1} debe dividir a Q_n . Puesto que $\text{gr}(Q_n) \leq n$, tenemos $Q_n = 0$, de donde $R_n = 0$.

⁽¹⁾ Puesto que P/Q es irreducible, necesariamente A/R y B/S son irreducibles. En efecto, si, por ejemplo, A/R no fuese irreducible, se tendría: $A/R = A_1/R_1$, con $\text{gr}(R_1) < \text{gr}(R)$. F admitiría entonces por representante $A_1/R_1 + B/S = (A_1S + BR_1)/R_1S$, con $\text{gr}(R_1S) < \text{gr}(Q)$, lo cual es absurdo.

b) *Existencia.* Hagamos

$$A = \sum_{k=0}^p a_k X^k \quad \text{y} \quad B = \sum_{k=0}^q b_k X^k.$$

Por hipótesis, $b_0 \neq 0$.

Razonemos por recurrencia sobre n . Para $n = 0$, es suficiente tomar

$$Q_0 = \frac{a_0}{b_0}, \quad R_0 = \frac{A - BQ_0}{X}.$$

Supongamos determinados Q_n y R_n , verificando (1). Puesto que $B(0) \neq 0$, existe una constante $\lambda_{n+1} = \frac{R_n(0)}{B(0)}$ y un polinomio S tales que

$$R_n = \lambda_{n+1} B + XS.$$

La relación (1) implica entonces:

$$A = BQ_n + \lambda_{n+1} X^{n+1} B + X^{n+2} R_n S.$$

Por lo tanto, es suficiente tomar $Q_{n+1} = Q_n + \lambda_{n+1} X^{n+1}$ y $R_{n+1} = R_n S$ para obtener (1) en el orden $n + 1$. c.q.d.

La disposición práctica de un cálculo de división sigue paso a paso el razonamiento constructivo anterior. Por ejemplo, he aquí la división de $A = 1 + 2X + X^3$ por $B = 1 + X + 2X^2$ hasta el orden 3:

$$\begin{array}{r} 1 + 2X + X^3 \\ X - 2X^2 + X^3 \\ - 3X^2 - X^3 \\ 2X^3 + 6X^4 \\ 4X^4 - 4X^5 \end{array} \quad \left| \begin{array}{l} 1 + X + 2X^2 \\ \hline 1 + X - 3X^2 + 2X^3 \end{array} \right.$$

El cociente del orden 3 es $1 + X - 3X^2 + 2X^3$, el resto es $4(1 - X)X^4$. En la mayoría de los casos sólo interesa el cociente, y el cálculo del último resto resulta inútil.

Vamos a aplicar VI.2.1 al cálculo de la parte polar relativa al polo a de la fracción $F = \frac{P(X)}{(X - a)^a Q_1(X)}$ (expresada en forma irreducible, con Q_1 no divisible por $X - a$, es decir $Q_1(a) \neq 0$).

Efectuemos el cambio de variable $X - a = U$; $F(X)$ se transforma en una fracción racional $G(U)$, siempre en forma irreducible:

$$G(U) = \frac{P(a + u)}{U^\alpha Q_1(a + U)};$$

0 es polo de orden α de $G(U)$. Hemos sido conducidos al caso en que $a = 0$.

TEOREMA VII.3.2

Dada la fracción irreducible

$$F = \frac{P(X)}{X^\alpha Q_1(X)}, \quad Q_1(0) \neq 0,$$

la parte polar de F relativa al polo 0 es la expresión

$$\frac{\lambda_0}{X^\alpha} + \frac{\lambda_1}{X^{\alpha-1}} + \dots + \frac{\lambda_{\alpha-1}}{X},$$

siempre que $\lambda_0 + \lambda_1 X + \dots + \lambda_{\alpha-1} X^{\alpha-1}$ sea el cociente en la división, según las potencias crecientes, de P por Q_1 hasta el orden $\alpha - 1$.

Demostración. Sea S el polinomio $\lambda_0 + \lambda_1 X + \dots + \lambda_{\alpha-1} X^{\alpha-1}$. Tenemos

$$P = SQ_1 + X^\alpha R_\alpha,$$

de donde

$$F = \frac{S}{X^\alpha} + \frac{R_\alpha}{Q_1} = \frac{\lambda_0}{X^\alpha} + \frac{\lambda_1}{X^{\alpha-1}} + \dots + \frac{\lambda_{\alpha-1}}{X} + \frac{R_\alpha}{Q_1}.$$

Según la última observación del § 2, $\frac{S}{X^\alpha}$ es la parte polar buscada. c.q.d.

Cuando $\alpha = 1$, la parte polar se expresa de forma simple. Se tiene:

$$F = \frac{P(X)}{Q(X)}, \quad Q(X) = (X - a) Q_1(X) \quad Q_1(a) \neq 0.$$

Según la fórmula de Taylor de orden 1 (válida para un cuerpo cualquiera) es $Q_1(a) = Q'_1(a)$. Por otro lado:

$$P(X) = \frac{P(a)}{Q_1(a)} Q_1(X) + (X - a) R_0, \quad \text{luego} \quad F = \frac{P(a)}{(X - a) Q'_1(a)} + \frac{R_0}{Q_1(X)}.$$

Podemos enunciar:

VII.3.3 *La parte polar relativa a un polo simple a de*

$$\left\| \quad F = \frac{P}{Q} \quad \text{es} \quad \frac{P(a)}{(X-a) Q'(a)} \right.$$

En característica 0, la fórmula de Taylor permitirá además obtener expresiones generales para las partes polares relativas a los polos de orden α .

§ VII.4 NOCIONES ACERCA DE LAS SERIES FORMALES

● En lo que sigue, K designa siempre un cuerpo conmutativo.

Una *serie formal con coeficientes en K* es una sucesión cualquiera $(a_n)_{n \geq 0}$, $a_n \in K$. A los (a_n) se les llama *coeficientes* de la serie. Se define la suma de dos series formales por:

$$(a_n) + (b_n) = (a_n + b_n)$$

y su producto:

$$(a_n) \times (b_n) = (c_n) \quad \text{por} \quad c_n = \sum_{p=0}^n a_p b_{n-p} \quad \text{para todo } n.$$

Con estas dos leyes, el conjunto (designado por $K[[X]]$) de las series formales es un anillo conmutativo. Evidentemente $K[X]$ se identifica con un subanillo de $K[[X]]$, a saber, el subanillo de las series que sólo tienen un número finito de coeficientes no nulos. El homomorfismo canónico $K \rightarrow K[X]$ define sobre $K[[X]]$ una estructura de K -álgebra. El elemento unidad de $K[[X]]$ es el de $K[X]$.

Por convenio, la serie formal (a_n) se designa por $\sum_{n \geq 0} a_n X^n$ o $\sum_{n=0}^{\infty} a_n X^n$. Esta notación es una extensión de la expresión de un polinomio con ayuda de la variable X .

El *orden* $\omega(S)$ de la serie formal $\sum_{n \geq 0} a_n X^n$ se define por medio de las fórmulas $\omega(0) = +\infty$, y, si $S \neq 0$, $\omega(S)$ es el menor entero n tal que $a_n \neq 0$.

Al orden de una serie formal también se le llama *valoración* de esta serie; esta noción generaliza la de valoración de un polinomio sobre un anillo (§ IV.1). Se verifican, sin dificultad, las propiedades:

$$\omega(S + T) \geq \inf(\omega(S), \omega(T)),$$

$$\omega(ST) = \omega(S) + \omega(T).$$

De estas fórmulas resulta, en particular, que $K[[X]]$ es un anillo íntegro.

TEOREMA VII.4.1

Sean $S = \sum_{m \geq 0} a_m X^m$ y $T = \sum_{m \geq 0} b_m X^m$ dos series formales, con $b_0 \neq 0$.
 Para todo entero n , existe un polinomio Q_n , y una serie formal R_n , definidos de forma única, tales que

$$S = TQ_n + X^{n+1} R_n, \quad \text{gr}(Q_n) \leq n.$$
Demostración

- a) La existencia se demuestra exactamente como en VII.3.1.
 b) *Unicidad.* Supongamos:

$$TQ_n + X^{n+1} R_n = 0, \quad \text{con } Q_n \neq 0 \quad \text{y} \quad \text{gr}(Q_n) \leq n.$$

Puesto que $b_0 \neq 0$, el orden de la serie formal TQ_n es $\leq n$, mientras que el orden de $X^{n+1} R_n$ es $\geq n+1$: hay, pues, contradicción. c.q.d.

COROLARIO

Para que la serie formal $S = \sum_{m \geq 0} a_m X^m$ sea invertible, es necesario y suficiente que $a_0 \neq 0$.

Demostración

- a) Si $ST = 1$, con $T = \sum_{n \geq 0} b_n X^n$, se tiene: $a_0 b_0 = 1$, de donde $a_0 \neq 0$.
 b) Si $a_0 \neq 0$, según VI.3.1, existe para todo n , un polinomio

$$Q_n = \sum_{k=0}^n \lambda_{n,k} X^k$$

y una serie formal R_n tales que:

$$(1) \quad SQ_n + X^{n+1} R_n = 1.$$

La unicidad del par (Q_n, R_n) nos permite afirmar que para $n < p$, $k \leq n$, se tiene $\lambda_{n,k} = \lambda_{p,k}$, pues de la relación $SQ_p + X^{p+1} R_p = 1$ se deduce:

$$S \left(\sum_{k=0}^n \lambda_{p,k} X^k \right) + X^{n+1} \left(X^{p-n} R_p + \sum_{k=n+1}^p \lambda_{p,k} X^{k-n} \right) = 1,$$

lo que prueba que el «polinomio truncado»

$$R'_p = \sum_{k=0}^n \lambda_{p,k} X^k$$

coincide con R_n .

Designemos por λ_k el valor común de los $\lambda_{n,k}$ para $n \geq k$.

Las relaciones (1) muestran que la serie $\sum_{k \geq 0} \lambda_k X^k$ es la inversa de S . c.q.d.

Como aplicación, observemos que *toda fracción racional, que no admite el 0 por polo, se puede desarrollar en serie formal.*

El teorema VII.4.1 da un procedimiento de cálculo de la serie entera que coincide con la fracción racional $\frac{P}{Q}$. *Es suficiente efectuar la división «indefinida» de P por Q según las potencias crecientes* ⁽¹⁾.

En lo que sigue, vamos a caracterizar las series formales que se pueden identificar con una fracción racional y dar otro método de cálculo de sus coeficientes. Comenzaremos por un caso particular:

VII.4.2 *El cuerpo K se supone de característica nula. La fracción racional*

$$\left\| \begin{array}{l} F_n = \frac{1}{(1-X)^n} \text{ admite el desarrollo en serie formal:} \\ (2) \quad S_n = \sum_{p \geq 0} \binom{n+p-1}{n-1} X^p. \end{array} \right.$$

Demostración. Por recurrencia sobre n . La propiedad resulta, cuando $n = 1$, de la fórmula $(1 + X + X^2 + \dots + X^n)(1 - X) = 1 - X^{n+1}$.

Suponemos que $F_{n-1} = S_{n-1}$ ($n \geq 2$).

Se tiene:

$$\begin{aligned} F_n &= \frac{F_{n-1}}{1-X} = F_{n-1} \left(\sum_{q \geq 0} X^q \right) = \left(\sum_{p \geq 0} \binom{n+p-2}{n-2} X^p \right) \cdot \left(\sum_{q \geq 0} X^q \right) \\ &= \sum_{m=0}^{\infty} \left(\sum_{0 \leq p \leq m} \binom{n+p-2}{n-2} \right) X^m. \end{aligned}$$

⁽¹⁾ Si se desea únicamente establecer el corolario de VII.4.1, la parte b) de la demostración se puede abreviar: buscando S^{-1} en la forma $S^{-1}(X) = \sum_{n \geq 0} b_n X^n$, se observa que los coeficientes b_n están unívoca-

mente determinados por las relaciones de recurrencia $a_0 b_0 = 1$ y (para $n \geq 1$) $a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0 = 0$. En efecto, puesto que $a_0 \neq 0$, b_0 está determinado, y para $n \geq 1$, b_n se expresa en función de $b_{n-1}, b_{n-2}, \dots, b_0$.

Luego, según la fórmula de Pascal:

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}, \quad \text{vemos que} \quad \sum_{0 \leq p \leq m} \binom{n+p-2}{n-2} = \binom{n+m-1}{n-1}.$$

de donde $F_n = \sum_{m=0}^{\infty} \binom{n+m-1}{n-1} X^m$. c.q.d.

Nota. Si K no es de característica nula, VII.4.2 subsiste, pero es preciso reemplazar en (2) los $\binom{n+p-1}{n-1}$ por los coeficientes $\binom{n+p-1}{n-1} \cdot 1$, en donde 1 designa al elemento unidad de K . Algunos de estos coeficientes son entonces nulos.

Se observará que VII.4.2 constituye una generalización de la fórmula del binomio, para un cuerpo cualquiera, y para los valores enteros negativos del exponente.

Es posible obtener de nuevo la fórmula (2) por derivación formal de $\frac{1}{1-X}$ (ver § 6).

Suponiendo que K sea algebraicamente cerrado, VI.4.2 nos da un nuevo procedimiento para calcular el desarrollo en serie de una fracción racional. En efecto, si escribimos la fracción en forma irreducible $F = \frac{P}{Q}$, con $Q(X) = \prod_{i=1}^r (X - a_i)^{\alpha_i}$, nos vemos conducidos, en virtud de VI.2.1, a escribir el desarrollo de términos de la forma $\frac{1}{(a - X)^{\alpha}}$, con $a \neq 0$. Pero, según VI.3.2,

$$\frac{1}{(a - X)^{\alpha}} = \frac{1}{a^{\alpha} \left(1 - \frac{X}{a}\right)^{\alpha}} = \frac{1}{a^{\alpha}} \sum_{p \geq 0} \binom{\alpha + p - 1}{\alpha - 1} \frac{X^p}{a^p}.$$

TEOREMA VII.4.3

Para que la serie formal $S = \sum_{p \geq 0} a_p X^p$ se pueda identificar con una fracción racional, es necesario y suficiente que existan dos enteros m, n ($n \geq m$) y constantes $\lambda_0, \lambda_1, \dots, \lambda_m$ con $\lambda_0 \neq 0$, tales que, para todo $p > n$ se verifique,

$$(3) \quad \lambda_0 a_p + \lambda_1 a_{p-1} + \dots + \lambda_m a_{p-m} = 0 \quad \text{con } p > n.$$

Demostración

a) Si $S = \frac{P(X)}{Q(X)}$, hacemos

$P(X) = \mu_0 + \mu_1 X + \dots + \mu_n X^n$, $Q(X) = \lambda_0 + \lambda_1 X + \dots + \lambda_m X^m$, $\lambda_0 \neq 0$, (conviene recordar que 0 *no es* polo de una fracción que se pueda identificar con una serie formal). Identificando los términos en X^p ($p > n$) en la relación

$$P(X) = Q(X) \cdot S(X),$$

se obtiene (3). Se observará que esta relación permite determinar los a_p por recurrencia.

b) Supongamos que se verifica (3). Entonces es inmediato que el producto de $S(X)$ por el polinomio $Q(X) = \lambda_0 + \lambda_1 X + \dots + \lambda_m X^m$ es un polinomio $P(X)$ de grado $\leq n$ (el coeficiente de X^p en SQ es nulo para $p > n$). Luego $S(X) = \frac{P(X)}{Q(X)}$

Observemos que en este caso $\frac{P(X)}{Q(X)}$ no es necesariamente la forma irreducible de la fracción $S(X)$. Se obtendrá la forma irreducible de $S(X)$ si el entero m es *mínimo* entre los que dan lugar a una relación de la forma (3). c.q.d.

Ejemplo

Desarrollemos en serie entera la fracción $F = \frac{C_0 + C_1 X}{1 - \alpha X - \beta X^2}$. Los coeficientes a_n de la serie buscada se obtienen por medio de las condiciones:

$$a_0 = C_0 \quad a_1 = \alpha C_0 + C_1,$$

y por la relación de recurrencia con dos términos:

$$(3) \quad a_n = \alpha a_{n-1} + \beta a_{n-2} \quad (n \geq 2)$$

Hemos visto en el capítulo V, § 5, que el término general a_n se escribe en la forma:

$$a_n = P_n a_1 + Q_n a_0,$$

con

$$P_2 = \alpha, \quad Q_2 = \beta, \quad Q_n = \beta P_{n-1} \quad (n \geq 3)$$

y

$$P_n = \sum_{0 \leq k \leq \frac{n-1}{2}} \binom{n-1-k}{k} \alpha^{n-1-2k} \beta^k.$$

Podemos, pues, escribir a continuación la fórmula completa:

$$\begin{aligned} \frac{C_0 + C_1 X}{1 - \alpha X - \beta X^2} = & C_0 + (\alpha C_0 + C_1) X + \\ & + \sum_{n \geq 2} \left\{ \left(\sum_{0 \leq k \leq \frac{n-1}{2}} \binom{n-1-k}{k} \alpha^{n-1-2k} \beta^k \right) (\alpha C_0 + C_1) \right. \\ & \left. + \left(\sum_{0 \leq k \leq \frac{n-2}{2}} \binom{n-2-k}{k} \alpha^{n-2-2k} \beta^{k+1} \right) C_0 \right\} X^n, \end{aligned}$$

expresión que tiene la ventaja de no presuponer nada acerca de las raíces de $1 - \alpha X - \beta X^2$.

Como aplicación numérica, tomemos $C_0 = 1$, $C_1 = 0$, $\alpha = \beta = 1$, con $K = \mathbf{R}$ o \mathbf{C} , de donde:

$$F = \frac{1}{1 - X - X^2}.$$

La fórmula anterior nos da:

$$F = 1 + X + \sum_{n \geq 2} X^n \left\{ \sum_{0 \leq k \leq \frac{n-1}{2}} \binom{n-1-k}{k} + \sum_{0 \leq k \leq \frac{n-2}{2}} \binom{n-2-k}{k} \right\}$$

También es posible expresar F introduciendo las raíces

$$r_1 = \frac{-1 - \sqrt{5}}{2}, \quad r_2 = \frac{-1 + \sqrt{5}}{2}$$

de $1 - X - X^2 = 0$; y observando que $r_1 r_2 = -1$:

$$\begin{aligned} F &= \frac{-1}{(r_1 - X)(r_2 - X)} = \frac{1}{r_1 - r_2} \left[\frac{1}{r_1 - X} - \frac{1}{r_2 - X} \right] \\ &= \frac{1}{r_1 - r_2} \left[\frac{1}{r_1} \sum_{n \geq 0} \frac{X^n}{r_1^n} - \frac{1}{r_2} \sum_{n \geq 0} \frac{X^n}{r_2^n} \right] \\ &= + \frac{\sqrt{5}}{5} \left[r_2 \sum_{n \geq 0} (-1)^n X^n r_2^n - r_1 \sum_{n \geq 0} (-1)^n X^n r_1^n \right] \\ &= \frac{\sqrt{5}}{5} \sum_{n \geq 0} (-1)^n X^n (r_1^{n+1} - r_2^{n+1}), \end{aligned}$$

fórmula que habríamos encontrado también *resolviendo* la relación de recurrencia (3) en este caso particular.

§ VII.5 EJEMPLOS DE CÁLCULOS PRÁCTICOS

A) Elementos de primera especie

(Caso en que el denominador se descompone en factores lineales en el cuerpo de base.)

$$1) \quad F = \frac{1}{X^4 + 1} \quad (\text{cuerpo } \mathbf{C}).$$

Los cuatro polos, simples, son las raíces cuartas de -1 , o sea

$$\omega_k = e^{i\frac{\pi}{4} + k\frac{\pi}{2}}, \quad 0 \leq k \leq 3.$$

Se tiene (en virtud de VII.3.3):

$$F = \sum \frac{A_k}{X - \omega_k}, \quad \text{con} \quad A_k = \frac{1}{4\omega_k^3} = -\frac{\omega_k}{4}.$$

Descompongamos ahora (siempre sobre \mathbf{C}) $F = \frac{1}{X^4 - 1}$. Los polos son ± 1 , $\pm i$. Por el mismo método, se obtiene:

$$F = \frac{A}{X-1} + \frac{B}{X+1} + \frac{C}{X-i} + \frac{D}{X+i} \quad A = \frac{1}{4}, \quad C = \frac{i}{4};$$

puesto que F es par, y debe permanecer invariante por el cambio $X \mapsto -X$, necesariamente $B = -A = -\frac{1}{4}$, $D = -C = -\frac{i}{4}$.

$$2) \quad F = \frac{P}{Q} = \frac{X^2 + 1}{X(X-1)^4(X^2-2)^2} \quad (\text{cuerpo } \mathbf{R}).$$

A priori, se tiene:

$$(1) \quad F = \frac{A}{X} + \frac{B_0}{(X-1)^4} + \frac{B_1}{(X-1)^3} + \frac{B_2}{(X-1)^2} + \frac{B_3}{X-1} + \frac{C_0}{(X-\sqrt{2})^2} \\ + \frac{C_1}{X-\sqrt{2}} + \frac{D_0}{(X+\sqrt{2})^2} + \frac{D_1}{X+\sqrt{2}}.$$

Se multiplica (1) por X , se reemplaza X por 0, de donde:

$$A = -\frac{1}{4}.$$

Se hace $X - 1 = Y$,

$$F = \frac{2 + 2Y + Y^2}{Y^4(1 - 3Y - 2Y^2 + 6Y^3 + Y^5)}.$$

Los coeficientes B_i se obtienen por medio de la división según las potencias crecientes hasta el orden 3 de $2 + 2Y + Y^2$ por $1 - 3Y - 2Y^2 + 6Y^3 + Y^5$. (Observemos que basta conocer los términos de grado ≤ 3 de este polinomio.)

Se obtiene:

$$B_0 = 2, \quad B_1 = 8, \quad B_2 = 29, \quad B_3 = 91.$$

Los coeficientes C_i, D_i se obtienen utilizando la misma técnica. Realicemos detalladamente el cálculo de C_0, C_1 . Se pone $X = \sqrt{2} + Y$, de donde:

$$F = \frac{3 + 2\sqrt{2}Y + Y^2}{(\sqrt{2} + Y)(\sqrt{2} - 1 + Y)^4(2\sqrt{2} + Y)^2 Y^2}.$$

Es suficiente calcular los términos de grado ≤ 1 del polinomio

$$(\sqrt{2} + Y)(\sqrt{2} - 1 + Y)^4(2\sqrt{2} + Y)^2.$$

o sea

$$8\sqrt{2}(\sqrt{2} - 1)^4 + [16(\sqrt{2} - 1)^4 + 32\sqrt{2}(\sqrt{2} - 1)^3]Y.$$

El cálculo de $(\sqrt{2} - 1)^3, (\sqrt{2} - 1)^4$ se realiza utilizando los restos de $(X - 1)^3, (X - 1)^4$ en la división por $X^2 - 2$, que son, respectivamente, $-7 + 5X, 17 - 12X$. Los términos de grado ≤ 1 buscados son entonces:

$$\begin{aligned} 8\sqrt{2}(17 - 12\sqrt{2}) + [16(17 - 12\sqrt{2}) + 32\sqrt{2}(-7 + 5\sqrt{2})]Y &= \\ &= -192 + 136\sqrt{2} + (592 - 416\sqrt{2})Y \\ Q_1 &= 8(-24 + 17\sqrt{2}) + 16(37 - 26\sqrt{2})Y. \end{aligned}$$

La división según las potencias crecientes hasta el orden 1 de $3 + 2\sqrt{2}Y + Y^2$ por Q_1 , proporciona C_0 y C_1 :

$$C_0 = \frac{3}{16}(24 + 17\sqrt{2}), \quad C_1 = \frac{1}{8}(365 + 258\sqrt{2}).$$

D_0 y D_1 se obtienen, evidentemente, reemplazando $\sqrt{2}$ por $-\sqrt{2}$ en C_0 y en C_1 (cálculos conjugados.)

$$3) F = \frac{1}{X^m(1-X)^n} \text{ (cuerpo } \mathbf{Q}).$$

Poniendo $1 - X = U$, se tiene:

$$F = \frac{1}{(1-U)^m U^n}.$$

Bastará, pues calcular la parte polar relativa al polo 0, lo que equivale a dividir, hasta el orden $m-1$, 1 por $(1-X)^n$. La manera más elegante de obtener el resultado consiste en escribir el desarrollo de $(1-X)^{-n}$ en serie formal:

$$(1-X)^{-n} = \sum_{k \geq 0} \binom{n+k-1}{n-1} X^k,$$

de donde resulta que la parte polar buscada es:

$$\mathcal{P}_0 = \sum_{k=0}^{m-1} \frac{\binom{n+k-1}{n-1}}{X^{m-k}}.$$

B) Elementos de segunda especie

(Caso en que el denominador contenga factores irreducibles de segundo grado sobre el cuerpo de base.)

$$1) F = \frac{1}{X^4 + 1} \text{ (cuerpo } \mathbf{R}).$$

Podríamos proceder al reagrupamiento de los términos conjugados, con la ayuda del ejemplo 1) de A). Es más rápido observar que

$$X^4 + 1 = (X^2 + 1)^2 - 2X^2 = (X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1),$$

de donde a priori:

$$(2) \quad F = \frac{AX + B}{X^2 - \sqrt{2}X + 1} + \frac{CX + D}{X^2 + \sqrt{2}X + 1}.$$

Puesto que F es par, $C = -A$ y $D = B$. Multiplicamos (2) por $X^2 - \sqrt{2}X + 1$, y reemplazamos X por una de las raíces, por ejemplo ζ , de este trinomio. Se obtiene:

$$\frac{1}{\zeta^2 + \sqrt{2}\zeta + 1} = A\zeta + B,$$

o sea, teniendo en cuenta que $\zeta^2 - \sqrt{2}\zeta + 1 = 0$,

$$(4A + 2\sqrt{2}B)\zeta - 2\sqrt{2}A - 1 = 0.$$

Dado que ζ es no real, se deduce:

$$2A + \sqrt{2}B = 0, \quad 2\sqrt{2}A + 1 = 0 \text{ de donde } A = -\frac{\sqrt{2}}{4}, \quad B = \frac{1}{2}.$$

De forma análoga se descompone, en \mathbf{R} , $F = \frac{1}{X^4 - 1}$. Según el ejemplo 1) de A), se tiene:

$$F = \frac{1}{4(X-1)} - \frac{1}{4(X+1)} + \frac{AX+B}{X^2+1};$$

multiplicando esta relación por $X^2 + 1$ y reemplazando X por i , se obtiene:

$$Ai + B = -\frac{1}{2}, \text{ de donde } A = 0, \quad B = -\frac{1}{2}.$$

($A = 0$ era previsible, en virtud de la paridad de F .)

2) $F = \frac{P}{Q} = \frac{X^2 + 1}{X(X-1)^4(X^2-2)^2}$ (cuerpo \mathbf{Q}). Los únicos polos en \mathbf{Q} son 0 y 1.

A priori, se tiene:

$$(2) \quad F = \frac{A}{X} + \frac{B_0}{(X-1)^4} + \frac{B_1}{(X-1)^3} + \frac{B_2}{(X-1)^2} + \frac{B_3}{X-1} + \frac{RX+S}{(X^2-2)^2} + \frac{TX+U}{X^2-2}.$$

En el ejemplo 1) de A), hemos calculado ya A, B_0, B_1, B_2, B_3 .

Multipliquemos (2) por $(X^2 - 2)^2$ y reemplacemos X por $\sqrt{2}$ (para efectuar los cálculos, se utilizan los restos módulo $(X^2 - 2)$ de todos los polinomios hallados). Se obtiene:

$$\frac{3}{\sqrt{2}(17 - 12\sqrt{2})} = R\sqrt{2} + S,$$

es decir: $34R - 24S - 3 + (17S - 24R)\sqrt{2} = 0.$

Puesto que $\sqrt{2} \notin \mathbf{Q}$, R y S vienen dados por las ecuaciones: $34R - 24S - 3 = 0$, $17S - 24R = 0$. Se deduce: $R = \frac{51}{2}$, $S = 36$.

(Se observará que en este cálculo, $\sqrt{2}$ ha desempeñado un papel meramente formal: el hecho de que sea real no ha servido para nada; sólo ha intervenido el hecho de que $\sqrt{2}$ sea una raíz de $X^2 - 2$ en un cierto supercuerpo de \mathbf{Q} .)

Para tener T y U , se procede «dando valores particulares» a X : $X = -1$ nos proporciona la primera relación:

$$-\frac{1}{8} = -A + \frac{B_0}{16} - \frac{B_1}{8} + \frac{B_2}{4} - \frac{B_3}{2} - R + S + T - U.$$

Multiplicando después (2) por X y «haciendo X infinito» (lo que equivale a comparar los grados de los numeradores tras haber reducido a denominador común en (2)), se obtiene la segunda relación:

$$0 = A + B_3 + T.$$

Estas dos relaciones nos dan, finalmente: $T = -\frac{365}{4}$, $U = -120$.

$$3) F = \frac{P}{Q} = \frac{2X^4 + 3X^3 + 4X^2 + X - 1}{(X-1)(X^2 + X + 1)^2} \text{ (cuerpo } \mathbf{R}).$$

La parte polar relativa al polo 1 es $\frac{1}{X-1}$; formamos $F - \frac{1}{X-1}$:

$$F - \frac{1}{X-1} = \frac{(X^2 - 1)(X^2 + X + 2)}{(X-1)(X^2 + X + 1)^2} = \frac{(X+1)(X^2 + X + 2)}{(X^2 + X + 1)^2} = G.$$

Utilizamos la proposición VII.2.2 tomando $P = X^2 + X + 1$:

$$G = \frac{X+1}{(X^2 + X + 1)^2} + \frac{X+1}{X^2 + X + 1},$$

de donde inmediatamente:

$$F = \frac{1}{X-1} + \frac{X+1}{(X^2 + X + 1)^2} + \frac{X+1}{X^2 + X + 1}.$$

$$4) F = \frac{1}{(X^{2n} - 1)^2} \text{ (cuerpo } \mathbf{R}).$$

En este caso lo más simple consiste en buscar ante todo la descomposición en \mathbf{C} :

$$F = \sum_{k=0}^{2n-1} \left[\frac{A_k}{(X - e^{ik\pi/n})^2} + \frac{B_k}{X - e^{ik\pi/n}} \right].$$

Hacemos $X = e^{-ip\pi/n} Y$, de donde

$$\begin{aligned} F &= \frac{1}{(Y^{2n} - 1)^2} = \sum_{k=0}^{2n-1} \left[\frac{A_k}{(Y - e^{ik\pi/n})^2} + \frac{B_k}{Y - e^{ik\pi/n}} \right] \\ &= \sum_{k=0}^{2n-1} \left[\frac{A_k}{(e^{-ip\pi/n} Y - e^{ik\pi/n})^2} + \frac{B_k}{e^{-ip\pi/n} Y - e^{ik\pi/n}} \right] \\ &= \sum_{k=0}^{2n-1} \left[\frac{A_k e^{2ip\pi/n}}{(Y - e^{i(k+p)\pi/n})^2} + \frac{B_k e^{ip\pi/n}}{Y - e^{i(k+p)\pi/n}} \right]. \end{aligned}$$

De ello se deduce que para todo entero p , $0 \leq p \leq 2n-1$,

$$A_p = A_0 e^{2ip\pi/n} \quad B_p = B_0 e^{ip\pi/n}.$$

El cálculo de A_0 y B_0 no presenta dificultad alguna; se hace $X = 1 + U$:

$$X^{2n} - 1 = U(2n + n(2n-1)U + \dots), \quad (1)$$

$$F = \frac{1}{U^2 [2n + n(2n-1)U + \dots]^2} = \frac{1}{U^2 [4n^2 + 4n^2(2n-1)U + \dots]}.$$

La división según las potencias crecientes de U por

$$4n^2(1 + (2n-1)U + \dots)$$

hasta el orden 1 conduce a los valores:

$$A_0 = \frac{1}{4n^2}, \quad B_0 = -\frac{2n-1}{4n^2}.$$

Resumiendo:

$$A_p = \frac{e^{2ip\pi/n}}{4n^2}, \quad B_p = -\frac{2n-1}{4n^2} e^{ip\pi/n}.$$

(1) Por convenio, los ... indican aquí series formales de orden superior al del último término escrito.

Para obtener la descomposición en \mathbf{R} , procedemos a reagrupar los términos conjugados; para $1 \leq p \leq n-1$:

$$\begin{aligned} \frac{A_p}{(X - e^{ip\pi/n})^2} + \frac{\bar{A}_p}{(X - e^{-ip\pi/n})^2} &= \frac{1}{4n^2} \left[\frac{e^{2ip\pi/n}}{(X - e^{ip\pi/n})^2} + \frac{e^{-2ip\pi/n}}{(X - e^{-ip\pi/n})^2} \right] = \\ &= \frac{1}{4n^2} \frac{2 \cos 2p \frac{\pi}{n} \cdot X^2 - 4 \cos p \frac{\pi}{n} \cdot X + 2}{\left(X^2 - 2 \cos p \frac{\pi}{n} \cdot X + 1 \right)^2} = \frac{1}{2n^2} \times \\ &\times \left[\frac{\cos 2p \frac{\pi}{n} \left(X^2 - 2 \cos p \frac{\pi}{n} X + 1 \right)}{\left(X^2 - 2 \cos p \frac{\pi}{n} X + 1 \right)^2} + 2 \sin^2 p \frac{\pi}{n} \frac{-2 \cos p \frac{\pi}{n} X + 1}{\left(X^2 - 2 \cos p \frac{\pi}{n} X + 1 \right)^2} \right], \\ \frac{B_p}{X - e^{ip\pi/n}} + \frac{B_p}{X - e^{-ip\pi/n}} &= -\frac{2n-1}{4n^2} \left[\frac{e^{ip\pi/n}}{X - e^{ip\pi/n}} + \frac{e^{-ip\pi/n}}{X - e^{-ip\pi/n}} \right] \\ &= -\frac{2n-1}{2n^2} \frac{\cos p \frac{\pi}{n} \cdot X - 1}{X^2 - 2 \cos p \frac{\pi}{n} X + 1}. \end{aligned}$$

La descomposición de F en el cuerpo \mathbf{R} es, pues:

$$\begin{aligned} F = \frac{1}{(X^{2n} - 1)^2} &= \frac{1}{4n^2 (X - 1)^2} - \frac{2n-1}{4n^2 (X - 1)} + \frac{1}{4n^2 (X + 1)^2} + \\ &+ \frac{2n-1}{4n^2 (X + 1)} + \sum_{p=1}^{n-1} \frac{\sin^2 p \frac{\pi}{n} \left(-2 \cos p \frac{\pi}{n} \cdot X + 1 \right)}{n^2 \left(X^2 - 2 \cos p \frac{\pi}{n} \cdot X + 1 \right)^2} \\ &+ \sum_{p=1}^{n-1} \frac{-(2n-1) \cos p \frac{\pi}{n} \cdot X + 2n-1 + \cos 2p \frac{\pi}{n}}{2n^2 \left(X^2 - 2 \cos p \frac{\pi}{n} \cdot X + 1 \right)}. \end{aligned}$$

C) Elementos cualesquiera

Ejemplo

Descomponer sobre el cuerpo \mathbf{Q} la fracción $F = \frac{1}{X(X-1)(X^3-2)}$.

$X^3 - 2$ es irreducible en \mathbf{Q} ; designemos por α su raíz real: los números $1, \alpha, \alpha^2$ son linealmente independientes en \mathbf{Q} (sin lo que $X^3 - 2$ admitiría un divisor de grado 1 ó 2 con coeficientes racionales).

A priori, tenemos:

$$F = \frac{A}{X} + \frac{B}{X-1} + \frac{CX^2 + DX + E}{X^3 - 2}.$$

Por el método habitual, se obtiene $A = \frac{1}{2}$, $B = -1$.

Multipliquemos la relación anterior por $X^3 - 2$, y reemplacemos X por α . Teniendo en cuenta que $\alpha^3 = 2$, se obtiene:

$$\frac{1}{\alpha(\alpha-1)} = C\alpha^2 + D\alpha + E;$$

$$1 = C\alpha^4 + (D-C)\alpha^3 + (E-D)\alpha^2 - E\alpha = (E-D)\alpha^2 + (2C-E)\alpha + 2(D-C);$$

$$(E-D)\alpha^2 + (2C-E)\alpha + 2(D-C) - 1 = 0.$$

La independencia de $1, \alpha, \alpha^2$ sobre \mathbf{Q} implica $E = D$, $2C = E$ y $2(D-C) - 1 = 0$, de donde $E = 1$, $D = 1$, y $C = \frac{1}{2}$.

§ VII.6 INTEGRACIÓN DE FRACCIONES RACIONALES

Derivación formal

Si K designa un cuerpo conmutativo cualquiera, sea $\frac{A}{B}$ un representante de la fracción $F \in K(X)$. Probemos que la fracción:

$$\frac{BA' - AB'}{B^2} \quad (\text{donde } P' \text{ designa la derivada formal del polinomio } P)$$

no depende del representante elegido. En efecto, si $\frac{A_1}{B_1}$ es un segundo representante de F , se tiene: $AB_1 = BA_1$. Debemos comprobar que

$$\frac{B_1 A'_1 - A_1 B'_1}{B_1^2} = \frac{BA' - AB'}{B^2},$$

o sea: $BB_1(BA'_1 - B_1 A') = B^2 A_1 B'_1 - B_1^2 AB'$;

En el segundo miembro de esta relación, reemplazamos BA_1 por AB_1 en el primer término, y AB_1 por BA_1 en el segundo; se obtiene:

$$BB_1(BA'_1 - B_1 A') = BB_1(AB'_1 - A_1 B').$$

Pero esta última relación se verifica, ya que, por derivación de $AB_1 = BA_1$, se tiene: $BA'_1 - B_1 A' = AB'_1 - A_1 B'$.

DEFINICIÓN VII.6.1

Se llama **derivada formal** de una fracción $F \in K(X)$, a la fracción $\frac{BA' - AB'}{B^2}$, en donde $\frac{A}{B}$ designa un representante cualquiera de F . Se la designa por F' , o $F'(X)$, o $\frac{dF}{dX}$ y se dice que F es una **primitiva (formal)** de F' .

Propiedades

- La aplicación $F \mapsto \frac{dF}{dX}$ prolonga la derivación formal definida ya en $K[X]$, y es K -lineal.
- Si $F \in K(X)$ y $G \in K(X)$, se tiene:

$$(1) \quad \frac{d}{dX}(FG) = \frac{dF}{dX} \cdot G + F \cdot \frac{dG}{dX}.$$

Para comprobar (1), se efectúa simplemente el cálculo, haciendo

$$F = \frac{P}{Q} \quad \text{y} \quad G = \frac{R}{S}, \quad P, Q, R, S \in K[X],$$

y observando que (1) es verdadera cuando F y G son polinomios.

VII.6.1 En el supuesto de que el cuerpo K sea de característica nula, para que
 \parallel la diferencia $F - G$ sea constante ($F, G \in K(X)$), es necesario y suficiente
 \parallel que $F' = G'$.

Demostración. Si $F = G + \text{Cte}$, se tiene $F' = G'$, pues la derivada de una constante es 0. Recíprocamente, supongamos que $F' = G'$, o sea $(F - G)' = 0$. Pongamos $F = G$ en forma irreducible: $F - G = \frac{P}{Q}$. La hipótesis significa que $QP' - PQ' = 0$, o bien: $QP' = PQ'$.

Puesto que P y Q son primos entre sí, el teorema de Gauss demuestra que Q divide a Q' y P divide a P' , lo cual implica: $Q = \text{Cte}$, $P = \text{Cte}$, luego $F - G = \text{Cte}$. c.q.d.

En este razonamiento se ha utilizado el hecho de que, si un polinomio divide a su derivada, es constante. Este resultado no es válido si K no es de característica nula. Por ejemplo, en $\mathbf{Z}/p\mathbf{Z}[X]$, el polinomio X^p tiene derivada 0.

Integración de fracciones racionales sobre \mathbf{C}

El problema es el siguiente: dada una función racional \tilde{F} , con coeficientes reales o complejos, hallar una primitiva de \tilde{F} .

Es claro que si la fracción racional F admite una primitiva formal G en $\mathbf{C}(X)$, la función racional \tilde{G} es una primitiva de \tilde{F} . Empezaremos, pues, estudiando el problema desde un punto de vista algebraico, buscando las fracciones racionales que admiten una primitiva racional. Para ello, resultará cómodo descomponer en elementos simples (fórmula 7 del § 2) las fracciones racionales consideradas y dar la siguiente definición:

DEFINICIÓN VII.6.2

} Sea a un polo de F de orden $\alpha \geq 1$, en donde F es una fracción racional sobre \mathbf{C} ; y sea

$$\mathcal{P}_a = \frac{A_1}{X - a} + \frac{A_2}{(X - a)^2} + \cdots + \frac{A_\alpha}{(X - a)^\alpha}$$

la parte polar de F relativa a este polo. Al número A_1 (coeficiente del término en $\frac{1}{X-a}$) se le llama **residuo** de F en a , y se designa por

$$A_1 = \text{Res}(F, a).$$

Una vez establecida esta definición, es fácil obtener el

TEOREMA VII.6.2

Para que una fracción racional con coeficientes en \mathbf{C} admita una primitiva racional, es necesario y suficiente que los residuos relativos a todos sus polos sean nulos.

Demostración

a) La condición es necesaria. Sea:

$$G = E + \sum_{i=1}^r \sum_{j=1}^{\alpha_i} \frac{A_{i,j}}{(X-a_i)^j}$$

(en donde E designa un polinomio y los $A_{i,j}$ designan constantes), una fracción racional cualquiera sobre \mathbf{C} , descompuesta en elementos simples. Su derivada

$$G' = E' - \sum_{i=1}^r \sum_{j=1}^{\alpha_i} j \frac{A_{i,j}}{(X-a_i)^{j+1}}$$

tiene todos sus residuos nulos.

b) La condición es suficiente.

Sea F una fracción racional que tenga todos sus residuos nulos. La parte entera de F , que es un polinomio, admite un polinomio por primitiva. Es suficiente, pues, probar que las partes polares de F poseen, cada una de ellas, una primitiva. Pero, por hipótesis, todas las partes polares de F son de la forma:

$$\mathcal{P}_a = \frac{A_2}{(X-a)^2} + \cdots + \frac{A_\alpha}{(X-a)^\alpha};$$

y la fracción \mathcal{P}_a admite como primitiva a la fracción:

$$G_a = -\frac{A_2}{X-a} - \frac{A_3}{2(X-a)^2} - \cdots - \frac{A_\alpha}{(\alpha-1)(X-a)^{\alpha-1}}.$$

Luego F admite una primitiva racional. c.q.d.

Ejemplos

La fracción racional

$$F = \frac{X^3 - 3X^2 - 1}{(X-1)^3} = 1 - \frac{3}{(X-1)^2} - \frac{3}{(X-1)^3}$$

admite como primitiva a

$$G = X + \frac{3}{X-1} + \frac{3}{2(X-1)^2} = \frac{2X^3 - 4X^2 + 8X - 3}{2(X-1)^2}.$$

Igualmente,

$$F = \frac{X^2 - 1}{(X^2 + 1)^2} = \frac{1}{2(X+i)^2} + \frac{1}{2(X-i)^2}$$

admite como primitiva a

$$G = -\frac{1}{2} \left[\frac{1}{X+i} + \frac{1}{X-i} \right] = \frac{-X}{X^2 + 1}.$$

Nota. En ciertos casos, se podrá determinar la primitiva de F sin tener necesidad de descomponer F en elementos simples. En particular, si F es de la forma:

$$F = \frac{P'}{P^n},$$

en donde P designa un polinomio, con $n \geq 2$, admite como primitiva a

$$G = \frac{1}{(1-n)P^n}.$$

Caso general

Si F es una fracción racional cuyos residuos no son todos nulos, es posible construir, por vía algebraica, una serie formal que sea una primitiva de la serie formal que representa a F , siempre que F no tenga al origen por polo. Pero esta serie carece de interés práctico por cuanto no se sabe «calcular» su suma. Abandonaremos, pues, momentáneamente el Álgebra; y, para terminar el cálculo de primitivas, nos limitaremos a las fracciones racionales con coeficientes reales. Es

evidente que bastará con determinar las primitivas de fracciones racionales de la forma

$$F = \sum_{i=1}^n \frac{A_i}{X - a_i},$$

en donde los A_i son constantes.

a) *Caso de un polo real*

Si a_i es real, la función $X \mapsto \frac{A_i}{X - a_i}$ admite como primitiva la función

$$X \mapsto A_i \operatorname{Log} |X - a_i|,$$

que no es racional.

b) *Caso de un polo complejo*

Si la fracción racional F tiene los coeficientes reales, y si admite un polo a no real, admite también \bar{a} como polo; y los residuos relativos a estos dos polos son imaginarios conjugados. Agrupando los términos relativos a estos polos, nos vemos conducidos a buscar una primitiva de una función de la forma

$$\Phi(X) = \frac{A}{X - a} + \frac{\bar{A}}{X - \bar{a}}.$$

Haciendo $A + \bar{A} = 2\alpha$, $a = p + iq$, $A\bar{a} + \bar{A}a = -\beta$, se tiene

$$\Phi(X) = \frac{2\alpha X + \beta}{X^2 - 2pX + q},$$

p, q, α, β son números reales tales que $p^2 - q < 0$.

Podemos escribir:

$$\Phi(X) = \frac{2\alpha(X - p) + \beta + 2\alpha p}{X^2 - 2pX + q} = \alpha \frac{(X^2 - 2pX + q)'}{X^2 - 2pX + q} + \frac{\beta + 2\alpha p}{X^2 - 2pX + q};$$

$\frac{(X^2 - 2pX + q)'}{X^2 - 2pX + q}$ admite como primitiva a $\operatorname{Log}(X^2 - 2pX + q)$ (siendo siempre > 0 la función que hay bajo el signo Log). Por otro lado,

$$\frac{1}{X^2 - 2pX + q} = \frac{1}{(X - p)^2 + q - p^2}$$

admite como primitiva a

$$\frac{1}{\sqrt{q - p^2}} \operatorname{arctg} \left(\frac{X - p}{\sqrt{q - p^2}} \right).$$

Luego Φ admite como primitiva a la función (no racional),

$$X \mapsto \alpha \operatorname{Log} (X^2 - 2 pX + q) + \frac{\beta + 2 \alpha p}{\sqrt{q - p^2}} \operatorname{arctg} \left(\frac{X - p}{\sqrt{q - p^2}} \right).$$

Ejemplo

$\frac{X}{X^2 + X + 1}$ admite como primitiva:

$$\frac{1}{2} \operatorname{Log} (X^2 + X + 1) + \frac{1}{\sqrt{3}} \operatorname{arctg} \left(\frac{2 X + 1}{\sqrt{3}} \right).$$

Capítulo VIII

Espacios vectoriales

Si bien no es indispensable, recomendamos al lector que no aborde este capítulo si no ha realizado la lectura del capítulo III, § 8.

§ VIII.1 GENERALIDADES

DEFINICIÓN VIII.1.1

Sea K un cuerpo cualquiera (no necesariamente conmutativo). Un espacio vectorial por la izquierda sobre K (o un K -espacio vectorial por la izquierda, o un espacio vectorial, cuando se trata de un solo cuerpo K) es un K -módulo por la izquierda.

Según se sigue del final del capítulo III, un espacio vectorial por la izquierda sobre K es, pues, un conjunto E provisto de dos leyes:

1) Una ley interna (designada aditivamente) que convierte a E en un grupo abeliano.

2) Una ley externa $(a, x) \mapsto a.x$ ($a \in K, x \in E$) de dominio K , tal que

$$\alpha(x_1 + x_2) = \alpha x_1 + \alpha x_2, \quad (\alpha_1 + \alpha_2)x = \alpha_1 x + \alpha_2 x \quad (\alpha_1, \alpha_2, \alpha \in K, x_1, x_2, x \in E)$$

y $\beta(ax) = (\beta a).x$ ($a \in K, \beta \in K, x \in E$); $1.x = x$ ($x \in E$), en donde 1 designa el elemento unidad de K .

— Recordemos (final del Cap. III) que se tienen las siguientes relaciones elementales:

$$\begin{aligned} 0_K \cdot x &= 0_E & (0_K \text{ elemento nulo de } K, 0_E \text{ elemento nulo de } E, x \in E), \\ \lambda \cdot 0_E &= 0_E & (\lambda \in K), \\ (-1) \cdot x &= -x & (x \in E). \end{aligned}$$

A los elementos de E se les llama ordinariamente, *vectores*; y, para distinguirlos, a los elementos de K se les llama entonces *escalares*.

A vectores x, y no nulos se les llama *colineales* si existe un $\lambda \in K$ tal que $y = \lambda x$.

— Sean E, F dos espacios vectoriales sobre el mismo cuerpo K . Una *aplicación lineal* (o *K-lineal*) de E en F , es un homomorfismo de espacios vectoriales sobre K , es decir, una aplicación f tal que

- L_1) para todo $x \in E$ y todo $y \in E$, $f(x + y) = f(x) + f(y)$;
 L_2) para todo $\lambda \in K$ y todo $x \in E$, $f(\lambda x) = \lambda f(x)$.

— Por definición, un *subespacio* del espacio vectorial E (sobre K) es un sub- K -módulo de E . Recordemos el teorema siguiente (cf. § III.8).

TEOREMA VIII.1.1

Para que una parte **no vacía** F de E sea un subespacio vectorial de E , es necesario y suficiente que, para todos $x_1, x_2, x \in F$ y todo $\lambda \in K$, se tenga

$$x_1 + x_2 \in F \quad \text{y} \quad \lambda \cdot x \in F.$$

En particular, el **núcleo** de una aplicación lineal de E en otro K -espacio vectorial F es un subespacio vectorial de E , y su **imagen** es un subespacio vectorial de F .

Espacio vectorial cociente

Sea E un K -espacio vectorial, F un subespacio y $p : E \rightarrow E/F$ el homomorfismo canónico del grupo aditivo E en el grupo cociente E/F .

Demostraremos que, para toda clase $X \in E/F$, y todos $x, x' \in X$, y todo $\lambda \in K$, se tiene: $p(\lambda x) = p(\lambda x')$. En efecto,

$$p(\lambda x) - p(\lambda x') = p(\lambda x - \lambda x') = p(\lambda(x - x')) ;$$

puesto que $x - x' \in F$ y puesto que F es un subespacio, se tiene: $\lambda(x - x') \in F$, de donde $p(\lambda(x - x')) = 0$. De esto resulta que $p(\lambda x)$ depende sólo de λ y de X .

Si hacemos $p(\lambda x) = \lambda.X$, se define una ley externa en E/F por medio de $(\lambda, X) \mapsto \lambda.X$, y se verifica que, provisto de esta segunda ley, E/F se convierte en un K -espacio vectorial.

Por definición, al K -espacio vectorial así construido se le llama *espacio vectorial cociente de E por F* , y se designa E/F . $p : E \rightarrow E/F$ es entonces una aplicación K -lineal.

TEOREMA VIII.1.2

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ p \downarrow & & \uparrow j \\ E/N & \xrightarrow{\bar{f}} & I \end{array}$$

E y F designan dos K -espacios vectoriales; sean $f : E \rightarrow F$ una aplicación lineal, N el núcleo de f e I su imagen, $p : E \rightarrow E/N$ la proyección canónica y $j : I \rightarrow F$ la inyección canónica, y $\bar{f} : E/N \rightarrow I$ el isomorfismo de grupos tal que:

$$f = j \circ \bar{f} \circ p.$$

Entonces \bar{f} es una aplicación lineal (luego, **un isomorfismo**) de E/N en el subespacio vectorial I de F .

Demostración (abreviada). La existencia y la unicidad de \bar{f} están garantizadas por los teoremas generales acerca de los grupos. Es suficiente ver que \bar{f} verifica (L_2) .

Sean $\lambda \in K$, $X \in E/N$, $x \in X$. Se tiene:

$$\begin{aligned} \bar{f}(\lambda X) &= \bar{f}(\lambda p(x)) = \bar{f}(p(\lambda x)) = (\bar{f} \circ p)(\lambda x) = f(\lambda x) = \lambda f(x) \\ &= \lambda.(\bar{f} \circ p)(x) = \lambda \bar{f}(X). \quad \text{c.q.d.} \end{aligned}$$

Suma directa de n espacios vectoriales (suma directa «externa»)

E_1, E_2, \dots, E_n ($n \in \mathbf{N}^*$) designan K -espacios vectoriales. Dotamos al producto cartesiano $E = E_1 \times E_2 \times \dots \times E_n$ de la ley de grupo abeliano, originada por las leyes de los E_i , y de la ley externa siguiente:

Si $x \in E$ y $\lambda \in K$, $x = (x_1, x_2, \dots, x_n)$ con $x_i \in E_i$,

$$(1) \quad \lambda.x = (\lambda x_1, \lambda x_2, \dots, \lambda x_n).$$

Provisto de estas leyes, E se convierte evidentemente en un K -espacio vectorial.

DEFINICIÓN VIII.1.2

Sean E_1, E_2, \dots, E_n espacios vectoriales. El producto cartesiano $E_1 \times E_2 \times \dots \times E_n$, provisto de la ley de grupo producto y de la ley externa definida por (1), es un espacio vectorial llamado **producto** de los E_i , o **suma directa** de los E_i .

La suma directa de los E_i se designa, a menudo, por $\bigoplus_{i=1}^n E_i$. Se puede designar también por $\prod_{i=1}^n E_i$. Así pues, las nociones de producto, y de suma directa, de un número finito de espacios vectoriales, coinciden. Se definen también estas nociones de producto y de suma directa en el caso de una familia infinita de espacios vectoriales, pero entonces no coinciden.

Sean E_1, E_2, \dots, E_n K -espacios vectoriales y $E = \bigoplus_{i=1}^n E_i$. Para todo entero $i \leq n$, sea $\psi_i : E_i \rightarrow E$ la aplicación tal que $\psi_i(t) = (a_i)_{1 \leq i \leq n}$, con $a_i = t$, $a_j = 0$ para $j \neq i$. ψ_i es una inyección K -lineal, llamada *canónica*, con ayuda de la cual se puede identificar E_i con un subespacio de E .

Sea $p_i : E \rightarrow E_i$ la i -ésima proyección. p_i es K -lineal y epiyectiva, su núcleo es el espacio vectorial $N_i = E_1 \times E_2 \times \dots \times E_{i-1} \times \{0\} \times E_{i+1} \times \dots \times E_n$. Según el teorema VIII.1.2, E/N_i es isomorfo a E_i .

Observemos, finalmente, que $p_i \circ \psi_i = I_i$, $p_i \circ \psi_j = 0$ para $i \neq j$. Además, si designamos por I a la aplicación idéntica de E por I_i a la de E_i , tenemos:

$$p_i \circ \psi_i = I_i, \quad \sum_{i=1}^n \psi_i \circ p_i = I.$$

— Sean F un K -espacio vectorial, y $f : F \rightarrow \bigoplus_{i=1}^n E_i$ una aplicación lineal. Con las notaciones anteriores, los $f_i = p_i \circ f$ son lineales.

Dar una aplicación lineal $f : F \rightarrow \bigoplus_{i=1}^n E_i$ equivale a dar n aplicaciones lineales $f_i = p_i \circ f$, llamadas *componentes* de f .

Igualmente, dar una aplicación lineal $g : \bigoplus_{i=1}^n E_i \rightarrow F$ equivale a dar n aplicaciones lineales $g_i = g \circ \psi_i$.

Suma de subespacios

Sean E un espacio vectorial, y F_1, F_2, \dots, F_n subespacios. La aplicación $\varphi : \bigoplus_{i=1}^n F_i \rightarrow E$ definida por

$$\varphi(x) = x_1 + x_2 + \dots + x_n \quad (x = (x_i)_{1 \leq i \leq n}; x_i \in F_i)$$

es K -lineal. Se le llama *aplicación canónica*. La imagen de φ es el subespacio vectorial $\sum_{i=1}^n F_i$ de E , suma de los F_i (recordemos que es el supremo de los F_i , en el conjunto de los subespacios de E ordenado por inclusión).

DEFINICIÓN VIII.1.3

Si F_1, F_2, \dots, F_n designan subespacios del K -espacio vectorial E . Se dice que la suma $\sum_{i=1}^n$ es **directa**, si la aplicación canónica de $\bigoplus_{i=1}^n F_i$ en E es inyectiva.
Se dice que **E es suma directa de los subespacios F_i** si esta aplicación canónica es biyectiva.

Decir que la suma de los subespacios F_i es directa significa que, si $x = x_1 + x_2 + \dots + x_n$, con $x_i \in F_i$ para todo i , la relación $x = 0$ implica $x_i = 0$ para todo i .
O también: si $x \in \sum F_i$, x se puede expresar de manera única en la forma $x = \sum_{i=1}^n x_i$ con $x_i \in F_i$ para todo i .

● Decir que E es suma directa de los F_i equivale a decir que *todo* $x \in E$ se puede expresar de una manera, y sólo una, en la forma $x = \sum_{i=1}^n x_i$, con $x_i \in F_i$ para todo i .

TEOREMA VIII.1.3

Si F_i ($1 \leq i \leq n$) designa subespacios vectoriales del K -espacio vectorial E , la suma de los F_i es directa si, y sólo si, se verifica la condición siguiente: para todo i ($1 \leq i \leq n$), se tiene:

$$F_i \cap \left(\sum_{j < i} F_j \right) = \{ 0 \}.$$

Demostración. La condición es necesaria, pues si la suma de los F_i es directa, y si $x \in F_i \cap \left(\sum_{j < i} F_j \right)$, se tiene: $x = \sum_{j < i} x_j$, con $x_j \in F_j$, de donde $x - \sum_{j < i} x_j = 0$, lo que implica $x = 0$ ya que $x \in F_i$.

La condición es suficiente: si suponemos que se verifica, y x es un elemento de $\sum F_i$, $x = x_1 + x_2 + \dots + x_n$, $x \in F_i$, para $1 \leq i \leq n$. Si $x = 0$, vemos que todos los x_i son nulos, pues si no lo fuesen, se podría definir el mayor de los enteros k tales que $x_k \neq 0$, y llamando i a este entero, se tendría: $x_i = - \sum_{j < i} x_j$ y $x_i \neq 0$, lo cual contradeciría la hipótesis. c.q.d.

● Caso particularmente importante

El espacio vectorial E es suma directa de *dos* subespacios F y G si, y sólo si, se tiene: $F \cap G = \{0\}$ y $F + G = E$.

Cuando esto ocurre, se dice que F y G son **suplementarios** (en E).

Sea E un espacio vectorial, suma directa de los subespacios $(F_i)_{1 \leq i \leq n}$. Todo $x \in E$ se expresa de manera única en la forma $x = \sum_{i=1}^n x_i$, con $x_i \in F_i$.

Pongamos $x_i = q_i(x)$. El estudio de las sumas directas externas realizado antes, muestra inmediatamente las siguientes propiedades: q_i es lineal, epiyectivo, de núcleo $\sum_{j \neq i} F_j$. Para $i \neq j$, $q_i \circ q_j = 0$, y $q_i^2 = q_i$. Finalmente $\sum_{i=1}^n q_i = I_E$ (en donde I_E es la aplicación idéntica de E). Resumiendo estas propiedades, se dice:

q_i es la proyección de E sobre F_i , paralelamente a $\sum_{j \neq i} F_j$.

En particular, si F y G son suplementarios en E , vemos (aplicando el teorema VIII.1.2 a las proyecciones de E sobre F y G) que G es isomorfo al espacio cociente E/F . Por lo tanto, todos los suplementarios de un subespacio dado F de E son isomorfos entre sí. (En el caso general, la *existencia* de tal suplementario no es evidente y resulta del axioma de la elección. Veremos una demostración elemental de esta existencia en el caso de los espacios de dimensión finita).

He aquí una aplicación importante del teorema VIII.1.2. Designemos por F y G dos subespacios de un K -espacio vectorial E , y sea $\varphi : F \oplus G \rightarrow E$ la aplicación canónica. El núcleo N de φ es el conjunto de los (x, y) ($x \in F, y \in G$) tales que $x + y = 0$. La imagen de φ es $F + G$. La aplicación $j : F \cap G \rightarrow N$, tal que $j(x) = (x, -x)$, es lineal e inyectiva. Es epiyectiva, pues si $x + y = 0$ ($x \in F, y \in G$), se tiene: $x = -y$, de donde $x \in G$ e $y \in F$, $x \in F \cap G$ e $y = -x$. Luego $F \cap G$ y N son isomorfos. Se deduce que el espacio vectorial $F + G$ es isomorfo al espacio cociente $(F \oplus G)/j(F \cap G)$:

VIII.1.4 Si F y G designan dos subespacios del K -espacio vectorial E , los espacios vectoriales $F + G$ y $(F \oplus G)/j(F \cap G)$ son isomorfos.

Restricción de escalares

Sean E un K -espacio vectorial, L un cuerpo, y ρ un isomorfismo de L en un subcuerpo de K . Dotemos a E de la ley externa siguiente (designada por $*$) de dominio L : si $b \in L$ y $x \in E$, es $b * x = \rho(b) \cdot x$.

Entonces E se convierte en un L -espacio vectorial, que se llama *espacio deducido de E por restricción a L de los escalares*. En particular, esto se aplicará cuando L sea un subcuerpo de K y $\rho : L \rightarrow K$ la inyección canónica.

El segundo problema: conociendo un L -espacio vectorial F , deducir un K -espacio vectorial, es mucho más delicado. Se trata muy elegantemente con la teoría del *producto tensorial*, pero sobrepasa los límites de nuestra obra.

§ VIII.2 CARACTERIZACIÓN DE LAS BASES DE UN ESPACIO VECTORIAL

K designa siempre un cuerpo cualquiera. En el capítulo III hemos visto lo que se entiende por familia generatriz, parte generatriz, familia libre, parte libre y base de un módulo. Por definición, una familia generatriz [resp. parte generatriz, familia libre, parte libre o base] de un espacio vectorial E sobre K es una familia generatriz [resp. parte generatriz, familia libre, parte libre o base] del K -módulo E . Por ejemplo, una *parte* S de E es *no libre*, o *ligada*, si existe una parte finita X de S , y una familia $(\lambda_x)_{x \in X}$ de elementos de K *no todos nulos*, tales que

$$\sum_{x \in X} \lambda_x \cdot x = 0.$$

En caso contrario, S es **libre**. A los elementos de una parte libre (resp. no libre) se les llama *linealmente independientes* (resp. *ligados* o *linealmente dependientes*).

Análogamente, una familia $(a_i)_{i \in I}$ de elementos de E recibe el nombre de **familia generatriz** si, para todo $x \in E$, existen escalares $\lambda_i \in K$ ($i \in I$) *casi todos nulos* (e.d. nulos a excepción de un número finito de ellos) tales que

$$x = \sum_{i \in I} \lambda_i a_i.$$

Una **base** de E es una familia a la vez libre y generatriz (cf. § III.8).

Vamos a establecer ahora un teorema que se aplica únicamente a los espacios vectoriales y que permite reconocer si una *parte* de un espacio vectorial constituye una base.

Recordemos (cf. Cap. III) que, por convenio, el conjunto vacío es una parte libre de E , y que es cómodo escribir

$$\sum_{x \in \emptyset} \lambda_x x = 0$$

a fin de no tener que analizar en las demostraciones demasiados casos particulares.

Ante todo estableceremos un lema:

VIII.2.1 Si x designa a un elemento no nulo del espacio vectorial E , $\{x\}$ es una
|| *parte libre* de E .

Demostración. Sea $\lambda \in K$ tal que $\lambda x = 0$. Se demuestra, por reducción al absurdo, que $\lambda = 0$. Si no lo fuese, λ sería invertible en K , y se tendría

$$\lambda^{-1} \cdot (\lambda x) = 0 = (\lambda^{-1} \lambda) \cdot x = 1 \cdot x = x,$$

lo cual es absurdo. c.q.d.

TEOREMA VIII.2.2

Sea \mathcal{B} una parte no vacía del K -espacio vectorial E . Las propiedades que siguen son equivalentes:

- a) \mathcal{B} es una base de E ;
- b) \mathcal{B} es una parte **generatriz** de E **minimal** respecto de la inclusión (es decir, \mathcal{B} es un elemento minimal en el conjunto de las partes generatrices de E , ordenado por inclusión);
- c) \mathcal{B} es una parte **libre** de E , **maximal** respecto de la inclusión.

Demostración. Bastará con establecer las implicaciones:

$$(a) \Rightarrow (b), \quad (b) \Rightarrow (a), \quad (a) \Rightarrow (c) \quad \text{y} \quad (c) \Rightarrow (a),$$

$(a) \Rightarrow (b)$. Si \mathcal{B} es una base, \mathcal{B} es una parte generatriz. Sean $b \in \mathcal{B}$ y $\mathcal{C} = \mathcal{B} \setminus \{b\}$. Si \mathcal{C} fuese una parte generatriz, existirían escalares $(\lambda_c)_{c \in \mathcal{C}}$ casi todos nulos, tales que $b = \sum_{c \in \mathcal{C}} \lambda_c c$. Escribiendo esta relación en la forma $1 \cdot b - \sum_{c \in \mathcal{C}} \lambda_c c = 0$, vemos que \mathcal{B} no es una parte libre. Luego \mathcal{C} no es una parte generatriz.

$(b) \Rightarrow (a)$. Sea \mathcal{B} una parte generatriz minimal de E . Para probar que \mathcal{B} es una base, es suficiente demostrar que \mathcal{B} es libre, pues, en caso contrario, existirían escalares $(\lambda_b)_{b \in \mathcal{B}}$ casi todos nulos, pero no todos nulos, tales que $\sum_{b \in \mathcal{B}} \lambda_b b = 0$. Sea ahora $b \in \mathcal{B}$ tal que $\lambda_b \neq 0$. La relación precedente implica $b = \sum_{c \in \mathcal{B}} (-\lambda_b^{-1} \lambda_c) c$, con $\mathcal{C} = \mathcal{B} \setminus \{b\}$.

Esta relación demuestra que \mathcal{C} sería una parte generatriz de E , estrictamente contenida en \mathcal{B} , contrariamente a la hipótesis.

$(a) \Rightarrow (c)$. Si \mathcal{B} es una base, entonces \mathcal{B} es una parte libre maximal. Sea $x \in E \setminus \mathcal{B}$. Existen escalares $(\lambda_b)_{b \in \mathcal{B}}$ casi todos nulos, tales que $x = \sum_{b \in \mathcal{B}} \lambda_b b$. Hacemos $\lambda_x = -1$ y $\mathcal{C} = \mathcal{B} \cup \{x\}$. La relación anterior se escribe $\sum_{c \in \mathcal{C}} \lambda_c c = 0$; y puesto que $\lambda_x = -1 \neq 0$, vemos que \mathcal{C} es no libre.

$(c) \Rightarrow (a)$. Si \mathcal{B} es una parte libre maximal, entonces \mathcal{B} es una parte generatriz. Sea $x \in E$. Si $x \in \mathcal{B}$, x es evidentemente combinación lineal de los elementos

de \mathcal{B} . Si $x \notin \mathcal{B}$, el conjunto $\mathcal{C} = \mathcal{B} \cup \{x\}$ es una parte no libre (puesto que \mathcal{B} es maximal). Existen entonces escalares $(\lambda_c)_{c \in \mathcal{C}}$ casi todos nulos, pero no todos nulos, tales que $\sum_{c \in \mathcal{C}} \lambda_c c = 0$; y se tiene $\lambda_x \neq 0$ (si no \mathcal{B} no sería libre). Esta relación implica, pues:

$$x = \sum_{b \in \mathcal{B}} (-\lambda_x^{-1} \lambda_b) b,$$

en otras palabras, x es combinación lineal de los elementos de \mathcal{B} . c.q.d.

Otro resultado fundamental relativo a las bases es el siguiente:

VIII.2.3 Sean E, F dos espacios vectoriales sobre un cuerpo K , $(e_i)_{i \in I}$ una base de E , $(a_i)_{i \in I}$ una familia cualquiera (con índices en I) de elementos de F . Entonces

a) Existe una aplicación lineal única

$$\varphi : E \rightarrow F$$

tal que:

$$\forall i \in I, \quad \varphi(e_i) = a_i.$$

b) Para que φ sea inyectiva (resp. epiyectiva, biyectiva), es necesario y suficiente que $(a_i)_{i \in I}$ sea una familia libre (resp. una familia generatriz, una base).

Demostración. El apartado b) es prácticamente evidente. Nos limitaremos a demostrar a): Si φ existe, sea $(\lambda_i)_{i \in I}$ una familia de escalares casi todos nulos. Se tiene necesariamente:

$$\varphi \left(\sum_{i \in I} \lambda_i e_i \right) = \sum_{i \in I} \lambda_i a_i.$$

De ahí la unicidad de φ .

Recíprocamente, a todo $x \in E$, asociamos la familia $(\lambda_i(x))_{i \in I}$ de sus coordenadas en la base $(e_i)_{i \in I}$. Estos $\lambda_i(x)$ son casi todos nulos, y la aplicación $\varphi : x \mapsto \sum_{i \in I} \lambda_i(x) a_i$ es lineal, y verifica: $\forall i \in I \quad \varphi(e_i) = a_i$. c.q.d.

La aplicación φ se ha obtenido «prolongando por linealidad» los $e_i \mapsto a_i$.

Nota. Sea K^I el K -espacio vectorial $\mathcal{F}(I, K)$ de las aplicaciones de I en K . El conjunto (designado por $K^{(I)}$, de las familias $(\lambda_i)_{i \in I}$ de escalares casi todos nulos

de K , forma un *subespacio vectorial* de K^I . El hecho de que $(e_i)_{i \in I}$ sea una base se interpreta por la propiedad de que la aplicación

$$\begin{aligned}\Gamma : E &\rightarrow K^{(I)} \\ x &\mapsto (\lambda_i(x))_{i \in I}\end{aligned}$$

es un *isomorfismo* de espacios vectoriales.

§ VIII.3 TEOREMA DE LA DIMENSIÓN FINITA

● En este § los espacios vectoriales considerados tienen por dominio de operadores un cuerpo K cualquiera.

En lo que sigue, si E es un espacio vectorial y A es una parte de E , el subespacio *engendrado* por A en E se designará por $\text{Vect}(A)$. Recordemos que es el conjunto de las *combinaciones lineales* de elementos de A (cf. § III.8, p. 131). Se tiene inmediatamente:

$$\begin{aligned}\text{Vect}(A \cup B) &= \text{Vect}(A) + \text{Vect}(B) \\ \text{Vect}(A \cap B) &\subset \text{Vect}(A) \cap \text{Vect}(B) \\ &\quad (\text{con el convenio: } \text{Vect}(\emptyset) = \{0\}), \\ \text{si } A &\subset B, \text{ Vect}(A) \subset \text{Vect}(B), \\ \text{Vect}(\text{Vect}(A)) &= \text{Vect}(A).\end{aligned}$$

DEFINICIÓN VIII.3.1

$\left\{ \begin{array}{l} \text{Un espacio vectorial } E \text{ es de } \mathbf{dimensión\ finita} \text{ si existe en } E \text{ una parte} \\ \text{generatriz } \mathbf{finita}; \text{ en caso contrario es de } \mathbf{dimensión\ infinita}. \end{array} \right.$

Propiedades inmediatas

a) Si E es de *dimensión finita*, el cociente de E por un subespacio cualquiera es de *dimensión finita*. (Pues la imagen de una parte generatriz de E por una aplicación lineal f es una parte generatriz de $f(E)$.)

b) Si E y F son de *dimensión finita*, $E \times F$ es de *dimensión finita*: pues si $\text{Vect}(S) = E$ y $\text{Vect}(T) = F$, se tiene $\text{Vect}(S \times T) = E \times F$.

Por el contrario, no es evidente que, si F es un subespacio de un espacio de *dimensión finita* E , F sea de *dimensión finita*. En efecto, si $\text{Vect}(S) = E$, y si no se verifica que $S \subset F$, en general $\text{Vect}(S \cap F)$ no es igual a F .

c) $\{0\}$ es un espacio de dimensión finita. Por definición, su dimensión es el número entero cero; a $\{0\}$ se le llama *espacio vectorial nulo*.

d) De toda parte generatriz Σ de un espacio vectorial de dimensión finita, se puede extraer una parte generatriz **finita** S .

En efecto, sea T una parte generatriz finita de E . Para todo $t \in T$, existen escalares $(\lambda_{t,s})_{s \in \Sigma}$, casi todos nulos, tales que $t = \sum_{s \in \Sigma} \lambda_{t,s} \cdot s$; designemos por S_t a la parte (finita) de Σ , formada por los $\lambda_{t,s}$ tales que $\lambda_{t,s} \neq 0$. Es evidente que la parte finita S de Σ definida por $S = \bigcup_{t \in T} S_t$ es generatriz. c.q.d.

TEOREMA VIII.3.1

Sean E un espacio de dimensión finita, L una parte libre de E y Σ una parte generatriz de E . Entonces L es finito, y existe una base finita B de E , tal que $L \subset B \subset (L \cup \Sigma)$. Luego: **toda parte libre L** (y, en particular, toda base) **de E es finita**; y (tomando $L = \emptyset$) **toda parte generatriz de E contiene una base de E** . Por consiguiente, **E admite al menos una base finita**.

Demostración. Designemos por S una parte finita de Σ que también sea generatriz (cf. propiedad d)). Sea \mathcal{S} el conjunto de las partes M de S tales que $L \cup M$ sea una parte libre de E . Es claro que $\emptyset \in \mathcal{S}$, y \mathcal{S} es no vacío. Por otra parte, puesto que S es finito, \mathcal{S} también lo es. Existe, pues, un conjunto $M_0 \in \mathcal{S}$, por lo menos, cuyo cardinal es *máximo*. Probemos que $S \subset \text{Vect}(L \cup M_0)$. En caso contrario, existiría un $s \in S$ tal que $s \notin \text{Vect}(L \cup M_0)$, y $L \cup M_0 \cup \{s\}$ sería una parte libre de E . Entonces se tendría:

$$(M_0 \cup \{s\}) \in \mathcal{S}, \quad \text{y} : \quad \text{card}(M_0 \cup \{s\}) = \text{card}(M_0) + 1.$$

lo cual es absurdo.

Puesto que $S \subset \text{Vect}(L \cup M_0)$, se tiene: $\text{Vect}(S) \subset \text{Vect}(L \cup M_0)$, o sea $E = \text{Vect}(L \cup M_0)$, lo que demuestra que $L \cup M_0$ es una parte generatriz de E . Puesto que $L \cup M_0$ es libre, es una base de E y, por lo tanto, una parte generatriz *minimal* de E (cf. VIII.2.2). Según la propiedad d) anterior, resulta que el conjunto $L \cup M_0$ es *finito*. A fortiori, L es *finito*, y la base $B = L \cup M_0$ verifica las relaciones enunciadas. c.q.d.

TEOREMA VIII.3.2 (llamado «de intercambio»)

Sea A una parte de un espacio vectorial E . Para todo $x \in E$ y todo $y \in E$, las relaciones $x \in \text{Vect}(A \cup \{y\})$ y $x \notin \text{Vect}(A)$ implican: $y \in \text{Vect}(A \cup \{x\})$.

Demostración. Por hipótesis, existen escalares casi todos nulos $(\lambda_a)_{a \in (A \cup \{y\})}$ tales que $x = \sum_{a \in (A \cup \{y\})} \lambda_a \cdot a$; y como $x \notin \text{Vect}(A)$, se tiene: $\lambda_y \neq 0$. La relación anterior implica pues:

$$y = \lambda_y^{-1} x - \sum_{a \in A} (\lambda_y^{-1} \cdot \lambda_a) \cdot a, \text{ de donde } y \in \text{Vect}(A \cup \{x\}) \text{ . c.q.d.}$$

Recordemos (III.8.4) que si $(a_i)_{i \in I}$ es una *familia libre* de un espacio vectorial E , la aplicación $i \mapsto a_i$, de I en E , es necesariamente inyectiva. Luego, si I es finito, el cardinal de I es igual al del conjunto asociado a la familia $(a_i)_{i \in I}$. Diremos que este cardinal es el *cardinal del conjunto de los (a_i)* , o el *número de elementos de (a_i)* .

TEOREMA VIII.3.3 (Teorema de la dimensión)

|| *En un espacio vectorial E de dimensión finita, todas las bases son finitas y tienen el mismo número de elementos.*

Demostración. Según las observaciones anteriores, podemos limitarnos a considerar las bases de E que son *partes de E* .

Si $E = \{0\}$, su única base es \emptyset , en virtud de los convenios realizados, y el teorema, en este caso, queda demostrado. Supongamos pues que $E \neq \{0\}$.

Según VIII.3.1, existe una parte finita B de E que es una base de E . Hagamos $n = \text{card}(B)$, y se tiene $n \geq 1$. Vemos en primer lugar que si C designa otra base de E , C es *finito* (cf. VIII.3.1). Para establecer que $\text{card}(C) = n$, escribiremos $p = \text{card}(B \cap C)$ ($0 \leq p \leq n$), y razonaremos por recurrencia sobre el entero $q = n - p$.

a) Si $q = 0$ se tiene $B \subset C$. Puesto que C es una parte libre, y B una parte libre maximal, se deduce $B = C$, de donde $\text{card}(C) = n$.

b) Supongamos demostrada la propiedad cuando $n - p = q$ ($q \leq n - 1$), y vamos a ver que es también verdadera cuando $n - p = q + 1$, lo que implicará que es verdadera para $n - p = 0, 1, 2, \dots, n$, es decir, cualquiera que sea p .

Consideremos entonces una base C de E , tal que el $\text{card}(B \cap C) = p = n - q - 1$. Esta relación, junto con las hipótesis, muestra que C no está contenido en B . Sea entonces $c \in C \setminus B$. El conjunto $C_1 = C \setminus \{c\}$ no es una parte generatriz de E y, por lo tanto, existe un $b \in B$ tal que $b \notin \text{Vect}(C_1)$. Como también se tiene $b \in \text{Vect}(C) = \text{Vect}(C_1 \cup \{c\})$, el teorema de intercambio (VIII.3.2) implica que $c \in \text{Vect}(C_1 \cup \{b\})$, de donde $\text{Vect}(C_1 \cup \{b\}) = \text{Vect}(C_1 \cup \{c\}) = E$. Pero el conjunto $C_2 = C_1 \cup \{b\}$ es una parte libre de E , luego C_2 es una base de E que tiene en común con B los p elementos de $B \cap C$ y el elemento b , que no pertenece a $B \cap C$. Las bases B y C_2 tienen, pues, $p + 1$ elementos comunes, y por hipótesis de recurrencia se tiene $\text{card}(C_2) = n$. Pero es claro que $\text{card}(C) = \text{card}(C_2)$, luego $\text{card}(C) = n$. c.q.d.

Notaciones

Antes de desarrollar las consecuencias de este teorema fundamental, precisaremos ciertas notaciones. Si I designa un conjunto finito, y si $(a_i)_{i \in I}$ es una base del espacio vectorial E , el conjunto asociado a esta familia se designará por $\{a_i\}_{i \in I}$. Por ejemplo, si $I = N_n^*$, $\{e_1, e_2, \dots, e_n\} = \{e_i\}_{1 \leq i \leq n}$, designará la *parte* de E formada por los (e_i) , mientras que $(e_1, \dots, e_n) = (e_i)_{1 \leq i \leq n}$; designará la *familia* (e_i) . Se observará que si se fija el conjunto $\{e_i\}_{1 \leq i \leq n}$, existen $n!$ familias de E , de la forma $(f_i)_{1 \leq i \leq n}$, que lo admiten por conjunto asociado. Dar una de estas familias equivale a dar una reordenación del conjunto N_n^* . A una familia finita se le llama, a veces, *sistema*.

DEFINICIÓN VIII.3.2

Sea E un K -espacio vectorial no nulo de dimensión finita. Al entero $n \geq 1$, número de elementos de una base cualquiera de E , se le llama **dimensión** del espacio vectorial E . Se designa por $\dim_K E$, o $\dim(E)$ (cuando sólo está en juego el cuerpo K , y no hay peligro de confusión). A la dimensión del espacio vectorial engendrado por una familia de elementos de un espacio vectorial se le llama **rango** de esta familia.

Vamos a desarrollar ahora las consecuencias del teorema de la dimensión.

VIII.3.4 Toda parte libre L de un espacio vectorial E de dimensión finita n tiene **a lo sumo** n elementos; toda parte generatriz S de E tiene **por lo menos** n elementos.

Demostración. En virtud de VIII.3.1 y de la nota que precede, existe una parte generatriz finita S_0 de E , tal que $S_0 \subset S$, y existe una base B de E tal que $L \subset B$; luego existe una base B' de E tal que $B' \subset S_0$ (según VIII.3.1 aplicado a $L = \emptyset$). Con la ayuda de VIII.3.3 se concluye la demostración, puesto que $\text{card}(B) = \text{card}(B') = n$. c.q.d.

VIII.3.5 En un espacio vectorial de dimensión finita n , toda parte libre L que posea n elementos es una base; toda parte generatriz S que posea n elementos es una base.

Demostración. Si volvemos a considerar el razonamiento de VIII.3.4, vemos que L está contenido en una base, y que S contiene una base. Basta pues con aplicar VIII.3.3. c.q.d.

VIII.3.6 *Todo subespacio F de un espacio E de dimensión finita es de dimensión finita y se tiene: $\dim(F) \leq \dim(E)$.*

Demostración. Toda parte libre de F tiene a lo sumo n elementos. Sea p el número máximo de elementos de una parte libre de F , y L una parte libre de F de cardinal p , luego se tiene $p \leq n$. Además, en F , L es una parte libre maximal, luego es una base (T. VIII.2.2). c.q.d.

Nota. Para todo $a \in E \setminus \{0\}$, el conjunto $Ka = \{\lambda a \mid \lambda \in K\}$ es un subespacio de dimensión 1 llamado **recta vectorial**.

VIII.3.7 *Si F es un subespacio de dimensión finita E y si $\dim(F) = \dim(E)$, se tiene $F = E$.*

Demostración. En virtud de VIII.3.2, toda base de F es una base de E . c.q.d.

Sea B una base de un espacio vectorial no nulo E , y consideremos una *partición finita* $\{C_1, C_2, \dots, C_n\}$ de B . En estas condiciones, E es suma directa de los subespacios $\text{Vect}(C_i)$. En efecto, es cierto que E es suma de los $\text{Vect}(C_i)$ y se comprueban sin dificultad las hipótesis de VIII.1.3. Como aplicación de esta propiedad, se tiene el:

TEOREMA VIII.3.8

|| *Todo subespacio F de un espacio vectorial de dimensión finita E admite un subespacio suplementario.*

Nota. Podemos suponer $E \neq F$ y $F \neq \{0\}$.

Demostración. Sea C una base de F , luego C es una parte libre de E , y existe una base B de F tal que $C \subset B$. Pongamos $D = B \setminus C$. Puesto que $\{C, D\}$ es una partición de B , $\text{Vect}(C) = F$ y $\text{Vect}(D)$ son suplementarios (cf. p. 285).

En el transcurso de la demostración de VIII.3.8 hemos utilizado el resultado siguiente, llamado «teorema de la base incompleta», que es un simple corolario de VIII.3.1 y de VIII.3.3:

● *Toda parte libre L de un espacio vectorial E de dimensión finita se puede completar hasta obtener una base de E adjuntando elementos convenientes de E , en número igual a $\dim(E) - \text{card}(L)$.*

VIII.3.9 *E y F designan espacios vectoriales de dimensión finita. Se tiene*

||
$$\dim(E \times F) = \dim(E) + \dim(F)$$

Demostración. Sean B una base de E y C una base de F . Pongamos $B_1 = B \times \{0\}$, $C_1 = \{0\} \times C$. $B_1 \cup C_1$ es una parte libre de E . Además, $B_1 \cup C_1$ es una parte generatriz de E , pues si $z = (x, y) \in E \times F$, y es $B = \{b_1, \dots, b_p\}$ y $C = \{c_1, \dots, c_q\}$, existen escalares $\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q$ tales que

$$x = \sum \lambda_i b_i \quad y \quad y = \sum \mu_j c_j, \text{ de donde } z = \sum \lambda_i (b_i, 0) + \sum \mu_j (0, c_j) . \text{ c.q.d.}$$

Por recurrencia, se deduce de VIII.3.9 que, si los E_i son de dimensión finita,

$$\dim \left(\bigoplus_{i=1}^n E_i \right) = \sum_{i=1}^n \dim (E_i) .$$

● COROLARIO. Si E es suma directa de los subespacios $(F_i)_{1 \leq i \leq n}$, se tiene también: $\dim (E) = \sum_{i=1}^n \dim (F_i)$.

TEOREMA VIII.3.10

|| Para todo subespacio F de un espacio vectorial E , de dimensión finita, el espacio E/F es de dimensión finita, y se tiene:

(1) $\dim (F) + \dim (E/F) = \dim (E)$.

Demostración. F admite un suplementario G , y el teorema resulta del hecho de que E/F y G sean isomorfos, y de VIII.3.9. ||

● COROLARIO

|| Sean E un espacio de dimensión finita y F un espacio vectorial.

|| Para toda aplicación lineal $f: E \rightarrow E$, si $\text{Ker } f$ e $\text{Im } f$ designan el núcleo y la imagen de f , respectivamente, se tiene:

(2) $\dim (\text{Ker } f) + \dim (\text{Im } f) = \dim (E)$.

Demostración. En virtud de VIII.1.2, $\text{Im } f$ es isomorfo a $E/\text{Ker } f$. Bastará aplicar (1). ||

Ejemplo fundamental de espacio de dimensión finita

En lo que sigue, δ_{ij} designa el símbolo de Kronecker: $\delta_{ij} = 0$ si $i \neq j$ y $\delta_{ii} = 1$. Consideremos el K -espacio vectorial K^n . Los elementos $(e_i)_{1 \leq i \leq n}$, en donde e_i designa al elemento de coordenadas $(\delta_{ij})_{1 \leq j \leq n}$, forman una base de K^n , llamada

base canónica. Así, K^n es de dimensión finita n , y es el *modelo* de todos los espacios de dimensión n sobre K . En efecto, sea E un K -espacio vectorial de dimensión n , y (b_1, \dots, b_n) una base de E ; la aplicación $f: K^n \rightarrow E$, tal que $f((\lambda_1, \dots, \lambda_n)) = \sum_{i=1}^n \lambda_i b_i$, es un *isomorfismo* de K^n en E . Sin embargo, *ninguno de estos isomorfismos es privilegiado*, y para determinar uno de ellos, basta con dar una *base ordenada* (b_1, \dots, b_n) de E .

Restricción de escalares

Sea K un subcuerpo del cuerpo L , y sea E un espacio vectorial de dimensión finita n sobre L . L es canónicamente un espacio vectorial sobre K , y suponemos que $\dim_K(L) = p$. En estas condiciones, *el espacio vectorial $E_{(K)}$, obtenido por restricción de los escalares a K (final del § 1), es de dimensión np .*

A fin de establecer esta propiedad, designemos por (b_1, \dots, b_n) una base de E sobre L , y por $(\gamma_1, \dots, \gamma_p)$ una base de L sobre K . Si $x \in E$, x se escribe:

$$x = \sum_{i=1}^n \lambda_i b_i, \quad \text{con } \lambda_i \in L.$$

Por hipótesis, existen escalares $\rho_{ij} \in K$ tales que

$$\lambda_i = \sum_{j=1}^p \rho_{ij} \gamma_j, \quad \text{de donde } x = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \rho_{ij} \gamma_j b_i.$$

Luego la familia $(\gamma_j b_i)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ es un sistema de generadores de $E_{(K)}$.

Supongamos por otra parte, que se tiene: $\sum_{i,j} \rho_{ij} \gamma_j b_i = 0$. Puesto que los (b_i) son libres sobre L , se deduce, para todo i , $\sum_j \rho_{ij} \gamma_j = 0$. Los γ_j son libres sobre K , y esto implica $\rho_{ij} = 0$ para todo i y todo j . Luego $(\gamma_j b_i)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ es una base de $E_{(K)}$.]

En particular, este resultado se aplicará cuando $K = \mathbf{R}$ y $L = \mathbf{C}$.

Se tiene que $\dim_{\mathbf{R}} \mathbf{C} = 2$, pues $(1, i)$ es evidentemente una base de \mathbf{C} sobre \mathbf{R} . Todo espacio vectorial de dimensión finita n sobre \mathbf{C} es, pues, un espacio vectorial de dimensión *par* $2n$ sobre \mathbf{R} .

Deberá tenerse mucho cuidado en no confundir la \mathbf{C} -linealidad y la \mathbf{R} -linealidad, cuando \mathbf{C} y \mathbf{R} estén en juego.

Codimensión

DEFINICIÓN VIII.3.9

Sean E un K -espacio vectorial y F un subespacio de E ; si E/F es de dimensión infinita, se dice que F es de **codimensión infinita**. Si E/F

$\left\{ \begin{array}{l} \text{es de dimensión finita, al entero } \dim(E/F) \text{ se le llama } \mathbf{codimensión} \text{ de } F, \\ \text{y se designa por } \operatorname{codim}_K(F) \text{ (o } \operatorname{codim}(F)), \text{ a los subespacios } F \text{ de } E \text{ tales} \\ \text{que } \operatorname{codim}(F) = 1 \text{ se les denomina } \mathbf{hiperplanos} \text{ de } E. \end{array} \right.$

Si G es suplementario de F , sabemos que E/F y G son isomorfos. Luego $\operatorname{codim}(F) = \dim(G)$.

En particular, si E es de dimensión finita, se deduce de (1):

$$(3) \quad \dim(F) + \operatorname{codim}(F) = \dim(E).$$

§ VIII.4 ESPACIOS VECTORIALES Y APLICACIONES LINEALES. RANGO DE UNA APLICACIÓN LINEAL

Sean E, F dos espacios vectoriales cualesquiera sobre el cuerpo K , y sea $f: E \rightarrow F$ una aplicación lineal. Se ve fácilmente que f transforma una parte generatriz de E en una parte generatriz de $f(E)$. Luego, si E es de dimensión finita, $f(E)$ es de dimensión finita.

DEFINICIÓN VIII.4.1

$\left\{ \begin{array}{l} \text{Sean } E \text{ un espacio vectorial de dimensión } \mathbf{finita}, F \text{ un espacio vectorial} \\ \text{y } f: E \rightarrow F \text{ una aplicación lineal. La dimensión del espacio imagen } f(E) \\ \text{se llama } \mathbf{rango} \text{ de } f, \text{ y se designa por } \operatorname{rg}(f). \end{array} \right.$

Las propiedades que siguen son inmediatas pero muy importantes:

$$1) \quad \operatorname{rg}(f) \leq \inf(\dim(E), \dim(F)).$$

Además, $\operatorname{rg}(f) = \dim(F)$ si, y sólo si, f es *epiyectiva*, y

$$\operatorname{rg}(f) = \dim(E)$$

si, y sólo si, f es *inyectiva* (cf. VIII.3.7 y VIII.3.10).

2) Sea de nuevo $f: E \rightarrow F$ una aplicación lineal, en donde E designa un espacio de dimensión finita n , y designemos por (e_1, \dots, e_n) una base de E . La familia $(f(e_1), \dots, f(e_n))$ es un sistema de generadores de la imagen I de f , por lo tanto, se tiene: $\dim(\operatorname{Im} f) = n$ si, y sólo si, $(f(e_1), \dots, f(e_n))$ es una familia libre de F .

Volvamos al caso de dos K -espacios vectoriales cualesquiera E, F y designemos por $\mathcal{L}_K(E, F)$ al conjunto de las aplicaciones lineales de E en F . Recordemos que es posible dotar a $\mathcal{L}_K(E, F)$ de una estructura de grupo abeliano. Para todo $f \in \mathcal{L}_K(E, F)$ y todo $g \in \mathcal{L}_K(E, F)$, $f + g$ es la aplicación:

$$x \mapsto f(x) + g(x) \quad (x \in E).$$

Para todo $f \in \mathcal{L}_K(E, F)$ y todo $\lambda \in K$, sea λf la aplicación

$$x \mapsto \lambda f(x) \quad (x \in E).$$

λf es un homomorfismo del grupo abeliano E en el grupo abeliano F , pero en general no es una aplicación lineal. En efecto, para $\mu \in K$, se tiene:

$$\lambda f(\mu x) = \lambda(\mu f(x)) = (\lambda\mu) f(x).$$

Decir que λf es lineal equivale a decir que

$$\lambda f(\mu x) = \mu((\lambda f)(x)) = \mu(\lambda f(x)) = (\mu\lambda) f(x),$$

por lo tanto, a decir que

$$(1) \quad (\lambda\mu) \cdot f(x) = (\mu\lambda) f(x) \quad \text{para } x \in E, \lambda \in K, \mu \in K,$$

lo cual, en general, no se verifica. Si $\lambda\mu = \mu\lambda$ para $\lambda \in K$ y $\mu \in K$, (1) se verifica, y es fácil ver:

VIII.4.1 Si K es conmutativo, la aplicación $(\lambda, f) \mapsto \lambda f$ de $K \times \mathcal{L}_K(E, F)$ en $\mathcal{L}_K(E, F)$ define, en el grupo $\mathcal{L}_K(E, F)$, una estructura de K -espacio vectorial.

TEOREMA VIII.4.2

Si el cuerpo de base K es **conmutativo** y si los espacios vectoriales E y F tienen dimensión finita, el espacio vectorial $\mathcal{L}_K(E, F)$ tiene dimensión finita, y se verifica:

$$\dim(\mathcal{L}_K(E, F)) = \dim(E) \times \dim(F).$$

Demostración. Sean (e_1, \dots, e_n) y (f_1, \dots, f_p) bases de E y de F , en donde $n = \dim(E)$ y $p = \dim(F)$.

Para todo elemento $x = \sum x_i e_i$ de E ($x_i \in K$) y toda aplicación lineal $u : E \rightarrow F$, se tiene:

$$(2) \quad u(x) = \sum u(x_i e_i) = \sum x_i u(e_i),$$

luego u está determinado de forma única por los $u(e_i)$, $1 \leq i \leq n$.

Pongamos

$$u(e_i) = \sum_{j=1}^p a_{ji} f_j \quad (a_{ji} \in K).$$

Utilizando la conmutatividad de K , y cambiando i y j , (2) implica:

$$(3) \quad u(x) = \sum_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} (x_i a_{ji}) f_j = \sum_{j=1}^p \left(\sum_{i=1}^n (a_{ji} x_i) \right) f_j.$$

Designemos por u_{ij} a la aplicación lineal de E en F tal que $u_{ij}(e_j) = f_i$ ($1 \leq i \leq p$) y $u_{ij}(e_k) = 0$ si $k \neq j$, (3) muestra que se verifica:

$$u = \sum_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} a_{ij} u_{ij}.$$

Luego la familia (u_{ij}) engendra $\mathcal{L}_K(E, F)$. Finalmente, la familia (u_{ij}) es libre, pues una relación de la forma

$$\sum \lambda_{ij} u_{ij} = 0 \quad (\lambda_{ij} \in K)$$

implica, por aplicación al elemento e_j : $\sum_{i=1}^p \lambda_{ij} f_i = 0$, y es $\lambda_{ij} = 0$ para $1 \leq i \leq p$ (puesto que $(f_i)_{1 \leq i \leq p}$ es libre), luego $\lambda_{ij} = 0$ para todo par (i, j) . Puesto que el número de los u_{ij} es np , el teorema está demostrado. \square

● Caso en que $E = F$

Sea de nuevo E un espacio vectorial sobre un cuerpo cualquiera K , y sea $\mathcal{L}_K(E)$ el grupo aditivo de los endomorfismos de E . Se define en $\mathcal{L}_K(E)$ una estructura de anillo estableciendo, para $u \in \mathcal{L}_K(E)$ y $v \in \mathcal{L}_K(E)$: $u \cdot v = u \circ v$ (compuesta de las aplicaciones u y v). Si $E \neq \{0\}$, $\mathcal{L}_K(E)$ es un anillo unífero, cuyo elemento unidad es la aplicación idéntica. En lo sucesivo, a este elemento unidad lo designaremos siempre por e .

Al grupo de los elementos invertibles del anillo $\mathcal{L}_K(E)$ se le llama *grupo lineal* de E , y se designa por $GL(E)$. Establecido esto, el teorema de la dimensión permite, cuando E es de dimensión finita, caracterizar los elementos de $GL(E)$:

TEOREMA VIII.4.3

Sea E un espacio vectorial de dimensión finita n sobre el cuerpo K , y sea u un endomorfismo de E . Las propiedades siguientes son equivalentes: (si $n \geq 1$)

a) u es invertible,	e) u es invertible por la derecha,
b) $\text{rg}(u) = n$,	f) u es invertible por la izquierda,
c) u es inyectiva,	g) u es regular por la derecha,
d) u es epiyectiva,	h) u es regular por la izquierda.

Demostración. Es inmediato que $a)$ implica las restantes propiedades $b)$ a $h)$. Además, $a)$, $b)$, $c)$ y $d)$ son equivalentes en virtud de VIII.3.5, VIII.3.7 y VIII.3.10. Puesto que $e)$ implica $g)$ y que $f)$ implica $h)$, bastará con establecer las implicaciones $g) \Rightarrow b)$ y $h) \Rightarrow b)$.

1) Para establecer $g) \Rightarrow b)$, probaremos que

$$(\operatorname{rg}(u) < n) \Rightarrow (\text{no } (g)).$$

Sea I la imagen de u . Puesto que $\operatorname{rg}(u) < n$, se tiene $\dim(I) < n$. Sea H un subespacio suplementario de I en E . Puesto que $\dim(I) + \dim(H) = n$ es $H \neq \{0\}$. Designemos por $p: E \rightarrow H$ a la proyección de E sobre H , paralelamente a I . p es epiyectiva, luego no nula ya que $H \neq \{0\}$ y evidentemente $p \circ u = 0$, luego u no es regular por la derecha.

2) Para establecer $h) \Rightarrow b)$, probaremos que: $(\operatorname{rg}(u) < n) \Rightarrow (\text{no } (h))$.

Sea N el núcleo de u . Puesto que $\operatorname{rg}(u) < n$, la relación:

$$\operatorname{rg}(u) + \dim(N) = n \quad \text{prueba que} \quad \dim(N) > 0, \quad \text{de donde: } N \neq \{0\}.$$

Sea H un subespacio suplementario de N en E , y sea $q: E \rightarrow N$ la proyección sobre N paralelamente a H . Puesto que q es epiyectivo y $N \neq \{0\}$, vemos que es $q \neq 0$. Luego es inmediato que $u \circ q = 0$, y u no es regular por la izquierda. c.q.d.

Cuando el cuerpo de base K es conmutativo (y suponiendo siempre que $n \geq 1$), $\mathcal{L}_K(E)$ es un espacio vectorial sobre K . La estructura de anillo de $\mathcal{L}_K(E)$, y esta estructura de espacio vectorial, hacen de $\mathcal{L}_K(E)$ una K -álgebra, como se comprueba inmediatamente. En este caso, si E es de dimensión finita n , $\mathcal{L}_K(E)$ es de dimensión finita n^2 (T. VIII.4.2).

Nota. En el caso en que el cuerpo de base K sea conmutativo, en virtud de lo que antecede, se tiene una demostración mucho más rápida de las implicaciones $g) \Rightarrow b)$ y $h) \Rightarrow b)$ del teorema VIII.4.3. Por ejemplo, probemos $g) \Rightarrow b)$ (suponiendo establecidas las otras propiedades).

Decir que u es regular por la derecha significa que la aplicación $\tau_u: f \rightarrow f \circ u$, de $\mathcal{L}_K(E)$ en sí mismo, es inyectiva. Puesto que τ_u es una aplicación lineal y $\mathcal{L}_K(E)$ es de dimensión finita, τ_u es biyectiva, y en particular, existe g tal que $g \circ u = e$, luego u es invertible por la izquierda. De esto se sigue que u es epiyectiva, luego biyectiva de E en E ya que E es de dimensión finita. c.q.d.

§ VIII.5 DUALIDAD

● En este § consideraremos únicamente espacios vectoriales sobre un cuerpo conmutativo K , y dotaremos a K de su estructura de espacio vectorial de dimensión 1 sobre K .

DEFINICIÓN VIII.5.1

Sea E un espacio vectorial. Al espacio vectorial $\mathcal{L}(E, K)$ de las aplicaciones lineales de E en K se le llama el espacio **dual** de E , y se designa por E^* . A los elementos de E^* se les denomina **formas lineales sobre E** .

Sean x un elemento de E y φ un elemento de E^* . Es cómodo designar el escalar $\varphi(x)$ por medio de $\langle x, \varphi \rangle$. La aplicación $(x, \varphi) \mapsto \langle x, \varphi \rangle$ de $E \times E^*$ en K verifica las siguientes propiedades:

$$(1) \quad \left. \begin{aligned} \langle x, \varphi_1 + \varphi_2 \rangle &= \langle x, \varphi_1 \rangle + \langle x, \varphi_2 \rangle \\ \langle x_1 + x_2, \varphi \rangle &= \langle x_1, \varphi \rangle + \langle x_2, \varphi \rangle \\ \lambda \langle x, \varphi \rangle &= \langle \lambda x, \varphi \rangle = \langle x, \lambda \varphi \rangle \end{aligned} \right\} \begin{aligned} &\lambda \in K, x, x_1, x_2 \in E, \\ &\varphi, \varphi_1, \varphi_2 \in E^*. \end{aligned}$$

Estas propiedades se resumen diciendo que $\langle x, \varphi \rangle$ es una forma *bilineal* de $E \times E^*$ en K . A esta forma bilineal se le llama forma bilineal *canónica* sobre $E \times E^*$.

DEFINICIÓN VIII.5.2

Sean E, F dos espacios vectoriales, E^* y F^* sus espacios duales. Sea $f: E \rightarrow F$ una aplicación lineal. Se llama **traspuesta de f** , y se designa por ${}^t f$, a la aplicación lineal:

${}^t f: F^* \rightarrow E^*$, tal que: ${}^t f(\varphi) = \varphi \circ f$ para todo $\varphi \in F^*$.

— El valor de ${}^t f(\varphi)$ sobre un elemento $x \in E$ es, pues, $\varphi(f(x)) = \langle f(x), \varphi \rangle$. Se dispone, pues, de la fórmula «mecánica» siguiente:

$$(2) \quad \langle x, {}^t f(\varphi) \rangle = \langle f(x), \varphi \rangle.$$

— Si E, F y G son espacios vectoriales y $f: E \rightarrow F$, $g: F \rightarrow G$ aplicaciones lineales, se tiene:

$${}^t(g \circ f) = {}^t f \circ {}^t g; \quad \text{además,} \quad {}^t(\text{id}_E) = (\text{id}_{E^*}).$$

— La aplicación $f \mapsto {}^t f$ de $\mathcal{L}_K(E, F)$ en $\mathcal{L}_K(F^*, E^*)$ es lineal.

DEFINICIÓN VIII.5.3

Sea A una parte no vacía del espacio vectorial E . Se llama **ortogonal de A en E^*** y se designa por A° , al conjunto de los $\varphi \in E^*$ tales que $\varphi(x) = 0$ para todo $x \in A$.

$\left\{ \begin{array}{l} \text{Se llama } \mathbf{ortogonal} \text{ en } E \text{ de una parte } A \text{ de } E^*, \text{ y se designa tam-} \\ \text{bién por } A^\circ, \text{ al conjunto de los } x \in E \text{ tales que, para todo } \varphi \in A, \text{ se veri-} \\ \text{fique } \varphi(x) = 0. \end{array} \right.$

Para toda parte A de E (resp. de E^*), A° es un subespacio de E^* (resp. de E). Se tienen las siguientes propiedades:

$$A \subset B \Rightarrow A^\circ \supset B^\circ \text{ de donde } A \subset A^{\circ\circ};$$

$$(A \cup B)^\circ = A^\circ \cap B^\circ \quad A, B \text{ partes de } E \text{ (o } E^*);$$

$$(\text{Vect}(A))^\circ = A^\circ \quad (\text{Vect}(A) : \text{subespacio engendrado por } A).$$

El teorema que sigue es esencial:

TEOREMA VIII.5.1

$\left\| \begin{array}{l} \text{Sean } E, F \text{ dos espacios vectoriales y } f: E \rightarrow F \text{ una aplicación lineal.} \\ \text{Entonces el núcleo de la traspuesta } {}^t f \text{ es igual al ortogonal de la ima-} \\ \text{gen de } f. \end{array} \right.$

Demostración. El núcleo de ${}^t f$ es el conjunto

$$\{ \psi \mid \psi \in F^* \text{ y } {}^t f(\psi) = 0 \}, \text{ o sea } \{ \psi \mid \psi \in F^* \text{ y } \psi \circ f = 0 \}.$$

Pero la relación $\psi \circ f = 0$ significa que $\langle f(x), \psi \rangle = 0$ para todo $x \in E$, y, por lo tanto, que ψ es ortogonal a $f(E)$. c.q.d.

Bidual

Sea E un espacio vectorial. Al dual de E^* se le llama el *bidual* de E , y se designa por E^{**} . Demostraremos ahora que existe una aplicación «natural» de E en E^{**} . Sea $x \in E$. La aplicación $\varphi \mapsto \langle x, \varphi \rangle$ de E^* en K es una forma lineal sobre E^* (fórmulas (1)). Designamos por \tilde{x} a esta aplicación. En virtud de las fórmulas (1), vemos que la aplicación:

$$J: E \rightarrow E^{**}$$

$$x \mapsto \tilde{x}$$

es lineal. A J se le llama aplicación canónica de E en su bidual. Se demuestra (con la ayuda del axioma de la elección) que J es siempre inyectiva. En general, J no es epiyectiva, pero lo es si E es de dimensión finita, como vamos a ver.

Dualidad y dimensión finita

Haciendo $F = K$ en el teorema VIII.4.2, se obtiene inmediatamente (puesto que por definición $\mathcal{L}_K(E, K) = E^*$).

TEOREMA VIII.5.2

|| Sea E un espacio vectorial de dimensión finita. El dual E^* de E es de dimensión finita, y $\dim(E^*) = \dim(E)$.

Sea $n = \dim(E)$, y (e_1, \dots, e_n) una base de E . Según la demostración del teorema VIII.4.2, una base de E^* está formada por los elementos e_i^* , en donde e_i^* es la forma lineal tal que

$$\langle e_j, e_i^* \rangle = \delta_{ij}$$

(δ_{ij} , símbolo de Kronecker, tal que $\delta_{ii} = 1$ y $\delta_{ij} = 0$ para $i \neq j$.)

● A la base (e_1^*, \dots, e_n^*) de E^* se le llama **base dual** de la base (e_1, \dots, e_n) .

Sea $x \in E$, con $x = \sum_{i=1}^n x_i e_i$ ($x_i \in K$). Se tiene:

$$e_i^*(x) = \langle x, e_i^* \rangle = \sum_{j=1}^n x_j \delta_{ij} = x_i.$$

Por esta razón, a la forma e_i^* se le llama, a veces, *la i -ésima forma coordenada* (relativa a la base (e_i)). En efecto: $e_i^*(x) = x_i$, lo que indica que el valor de e_i^* sobre el elemento $x \in E$ es igual a la i -ésima coordenada de x .

TEOREMA VIII.5.3

|| Si E es un espacio vectorial de dimensión finita, la aplicación canónica de E en su bidual es un isomorfismo.

Demostración. Según el teorema VIII.5.2,

$$\dim(E^{**}) = \dim(E^*) = \dim(E).$$

Bastará, pues, demostrar que J es inyectiva, por lo tanto que la relación $J(x) = 0$ implica $x = 0$, o que la relación $x \neq 0$ implica $J(x) \neq 0$.

Si $x \neq 0$, existe una base (e_1, \dots, e_n) de E tal que $e_1 = x$. Sea (e_1^*, \dots, e_n^*) la base dual de esta base. Por definición, se tiene:

$$\langle e_1^*, J(x) \rangle = \langle x, e_1^* \rangle,$$

de donde

$$\langle e_1^*, J(x) \rangle = \langle e_1, e_1^* \rangle = 1, \quad \text{luego } J(x) \neq 0 \text{ . c.q.d.}$$

Con la ayuda del isomorfismo $J: E \rightarrow E^{**}$, podemos identificar los espacios E y E^{**} . Cuando se realiza esta identificación, toda propiedad de dualidad demostrada para el par (E, E^*) nos da una propiedad análoga para el par (E^*, E) (puesto que, aplicada a E^* , esta propiedad es una propiedad del par (E^*, E^{**})).

Dicho de otra manera, *cuando E es de dimensión finita, existe una simetría total entre los papeles de E y E^** . Por ejemplo, si (e_1, \dots, e_n) es una base de E , esta base se identifica con la base dual de la base dual (e_1^*, \dots, e_n^*) de E^* definida antes del teorema VIII.5.3.

TEOREMA VIII.5.4

|| Si E es un espacio vectorial de dimensión finita, H es un subespacio de E y H° es su ortogonal, se tiene

$$\dim(H) + \dim(H^\circ) = \dim(E) .$$

Demostración. Sea $n = \dim(E)$, $p = \dim(H)$, y sea (e_1, \dots, e_n) una base de E tal que (e_1, \dots, e_p) sea una base de H . H° es el conjunto de las formas lineales φ sobre E tales que, para todo i comprendido entre 1 y p , se tiene $\langle e_i, \varphi \rangle = 0$, o sea $\varphi(e_i) = 0$. Descompongamos φ en la base dual (e_1^*, \dots, e_n^*) :

$$\varphi = \sum_{i=1}^n \lambda_i e_i^* \quad (\lambda_i \in K) .$$

$$\langle e_i, \varphi \rangle = 0 \text{ se escribe } \sum_{j=1}^n \lambda_j \langle e_i, e_j^* \rangle = 0; \text{ es decir: } \lambda_i = 0 .$$

Luego H° es el conjunto de los $\varphi \in E^*$ tales que $\lambda_i = 0$ para $1 \leq i \leq p$. Dicho con otras palabras, H° es el subespacio de dimensión $n - p$ de E^* engendrado por $(e_{p+1}^*, \dots, e_n^*)$. c.q.d.

Según la nota que precede al teorema VIII.5.4, se deduce del teorema VIII.5.4 que, si H es un subespacio de E^* , se tiene:

$$\dim(H) + \dim(H^\circ) = \dim(E) ,$$

siendo H° el ortogonal de H en $E^{**} = E$.

COROLARIO

|| Sea H un subespacio del espacio vectorial de dimensión finita E . Se tiene:
 $H^{\circ\circ} = H$.

|| Igualmente, si H es un subespacio de E^* , se tiene $H^{\circ\circ} = H$ (por la dualidad entre E y E^*).

Basta con demostrar la primera afirmación. Sabemos que $H^{\circ\circ} \supset H$. Además,

$$\dim(H) + \dim(H^\circ) = \dim(H^\circ) + \dim(H^{\circ\circ}) = \dim(E),$$

de donde $\dim(H) = \dim(H^{\circ\circ})$. Luego $H = H^{\circ\circ}$. c.q.d.

Este corolario no es trivial y tiene consecuencias importantes. Lo aplicaremos al estudio del núcleo de una forma lineal no nula.

Recordemos, ante todo, la siguiente:

DEFINICIÓN VIII.5.4

§ En un espacio vectorial E de dimensión finita n , a todo subespacio F de
§ codimensión 1 se le llama **hiperplano vectorial** (cf. § 6).

En virtud de la relación (3) del § VIII.3, vemos que los hiperplanos vectoriales de E son, precisamente, los subespacios de dimensión $n - 1$ de E .

Sea, entonces, φ un elemento de E^* , no nulo, y sea $D = K.\varphi$ el subespacio de dimensión 1 engendrado por φ en E^* . El ortogonal D° de D es, evidentemente, el núcleo de φ , y es también el núcleo de toda forma $\psi \in D \setminus \{0\}$. Según el teorema VIII.5.4, se tiene: $\dim(D^\circ) = \dim(E) - 1$. Luego D° es un hiperplano vectorial de E . El corolario de VIII.5.4 demuestra inmediatamente que $D^{\circ\circ} = D$, lo que significa que *toda función lineal sobre E , nula sobre D° , es un múltiplo de φ* .

Con otras palabras, *todo hiperplano vectorial de E admite, a lo sumo, una ecuación de la forma $\varphi = 0$, en donde $\varphi \in E^* \setminus \{0\}$, con la condición de considerar como idénticas las ecuaciones $\lambda\varphi = 0$ ($\lambda \in K^*$). O también:*

VIII.5.5 Sean φ, ψ dos formas lineales sobre E , y sea $\varphi \neq 0$. Si la relación

$$\left\| \begin{array}{l} \varphi(x) = 0 \text{ implica la relación } \psi(x) = 0, \text{ existe un escalar } \lambda \in K \text{ tal que} \\ \psi = \lambda\varphi. \end{array} \right.$$

Tomemos ahora un hiperplano vectorial cualquiera H de E . El teorema VIII.5.4 nos muestra que: $\dim(H) + \dim(H^\circ) = \dim(E)$, de donde: $\dim(H^\circ) = 1$. Luego existen formas lineales φ , todas proporcionales, cuyo núcleo es exactamente H ; es decir, H admite una ecuación. En resumen, si designamos por $\mathcal{G}_{n-1}(E)$ al conjunto de los hiperplanos vectoriales de E , y por $\mathcal{G}_1(E^*)$ al conjunto de los subespacios de dimensión 1 (o rectas) de E^* , la aplicación $H \mapsto H^\circ$ es una biyección de $\mathcal{G}_{n-1}(E)$ en $\mathcal{G}_1(E^*)$.

Generalizando, designemos por $\mathcal{G}_p(E)$ al conjunto de los subespacios de dimensión p de un espacio vectorial E . Razonando como anteriormente, se deduce del teorema VIII.5.4 y de su corolario el resultado siguiente:

VIII.5.6 Cuando el espacio vectorial E es de dimensión finita, la aplicación
 $\parallel F \mapsto F^\circ$ es una biyección de $\mathcal{G}_p(E)$ en $\mathcal{G}_{n-p}(E^*)$ ($n = \dim(E)$).

Apliquemos ahora el corolario de VIII.5.4 a la situación siguiente: el espacio E es de dimensión n , H es el subespacio engendrado en E^* por las formas lineales $\varphi_1, \dots, \varphi_p$; H° es el ortogonal de $\{\varphi_1, \dots, \varphi_p\}$. Si H_i designa al hiperplano de ecuación $\varphi_i = 0$, se tiene:

$$H^\circ = \bigcap_{i=1}^p H_i.$$

La propiedad $H^{\circ\circ} = H$, se puede enunciar entonces como sigue:

VIII.5.7 Si H_1, \dots, H_p designan los hiperplanos vectoriales de un espacio de dimensión finita, de ecuaciones respectivas $\varphi_1 = 0, \varphi_2 = 0, \dots, \varphi_p = 0$ ($\varphi_k \in E^* \setminus \{0\}$), toda forma lineal nula sobre $\bigcap_{i=1}^p H_i$, es una combinación lineal de $\varphi_1, \varphi_2, \dots, \varphi_p$.

Con la ayuda de estos resultados, se puede fundamentar rigurosamente la teoría geométrica de los haces (o de las redes) de planos del espacio. Consideremos, en \mathbb{R}^3 , dos planos P_1 y P_2 , de ecuaciones $\varphi_1 = 0$ y $\varphi_2 = 0$, entonces todo plano de ecuación $\lambda_1 \varphi_1 + \lambda_2 \varphi_2 = 0$ pasa por la recta $\Delta = P_1 \cap P_2$. La teoría de la dualidad nos prueba que el recíproco es verdadero, a saber, que todo plano que pasa por Δ tiene una ecuación de la forma $\lambda_1 \varphi_1 + \lambda_2 \varphi_2 = 0$.

Veremos más adelante numerosas aplicaciones del teorema VIII.5.4 y de su corolario.

Para terminar damos una consecuencia importante del teorema VIII.5.4.

TEOREMA VIII.5.8

\parallel Sean E, F dos espacios vectoriales de dimensión finita, y $f: E \rightarrow F$ una aplicación lineal. Si ${}^t f$ es la traspuesta de f , se tiene:

$$\text{rg}(f) = \text{rg}({}^t f).$$

Demostración. El núcleo N de ${}^t f$ es $(f(E))^\circ$ (T. VIII.5.1).

Por una parte se tiene:

$$\dim(f(E)) + \dim(f(E))^\circ = \dim(E),$$

$$\dim(f(E)) = \text{rg}(f),$$

y por otra:

$$\dim(N) + \operatorname{rg}({}^t f) = \dim(E).$$

De donde $\operatorname{rg}(f) = \operatorname{rg}({}^t f)$. c.q.d.

* Los resultados que preceden se extienden en parte a los espacios vectoriales denominados de **dimensión infinita** (e.d. que no son de dimensión finita).

TEOREMA VIII.5.9

Sean E un espacio vectorial cualquiera y H un subespacio de **dimensión finita** del dual E^* . Para la dualidad entre E y E^* , se tiene entonces:

$$H^{\circ\circ} = H.$$

Demostración. Razonaremos por recurrencia sobre $p = \dim(H)$.

a) Cuando $p = 1$, H admite por base a $\{\varphi\}$, en donde $\varphi \in E^* \setminus \{0\}$, y H° es el núcleo de φ . El rango de la aplicación $\varphi : E \rightarrow K$ es 1, φ es epiyectiva, y el cociente E/H° es isomorfo al K -espacio vectorial K . Existe una aplicación lineal $\bar{\varphi} : E/H^\circ \rightarrow K$, única, tal que el diagrama que sigue es conmutativo (descomposición canónica de φ). $\bar{\varphi}$ es un isomorfismo.

$$\begin{array}{ccc} E & \xrightarrow{\text{(apl. can.)}} & E/H^\circ \\ & \searrow \varphi & \downarrow \bar{\varphi} \\ & & K \end{array}$$

Pero para toda forma lineal ψ sobre E , nula sobre H° , existe $\bar{\psi} : E/H^\circ \rightarrow K$ única tal que $\bar{\psi} \circ p = \psi$ ⁽¹⁾; y para todo $\theta \in (E/H^\circ)^*$, $\theta \circ p \in E^*$. Así $\psi \rightarrow \bar{\psi}$ es una biyección de $H^{\circ\circ}$ en el dual de E/H° . Dado que E/H° es de dimensión 1, y que la aplicación $\psi \mapsto \bar{\psi}$ es lineal, se tiene que $H^{\circ\circ}$ es de dimensión 1, puesto que $\dim((E/H^\circ)^*) = \dim(E/H^\circ) = 1$.

b) Supongamos el teorema verdadero cuando $\dim(H) = p$, y demostrémoslo para $\dim(H) = p + 1$. Sea $(\varphi_1, \dots, \varphi_p, \varphi_{p+1})$ una base de H , y para todo i sea H_i el núcleo de φ_i . Sea, finalmente, $\varphi \in H^{\circ\circ}$. Designemos por ψ_i a la restricción de φ_i a H_{p+1} ($1 \leq i \leq p$), y por ψ a la restricción de φ a H_{p+1} .

Puesto que φ se anula sobre $H_1 \cap H_2 \cap \dots \cap H_{p+1}$, la función ψ se anula sobre

$$(H_1 \cap \dots \cap H_p) \cap H_{p+1},$$

es decir, sobre la intersección de los núcleos de los ψ_i . Según la hipótesis de recurrencia, ψ es una combinación lineal $\sum_{i=1}^p \lambda_i \psi_i$ de los ψ_i . En virtud de la parte a) de la demostración, la forma lineal

$$\theta = \varphi - \sum_{i=1}^p \lambda_i \varphi_i,$$

(1) «Propiedad universal del cociente» (cf. § 6, ejemplo 1).

(que se anula sobre H_{p+1}), es igual a $\lambda_{p+1}\varphi_{p+1}$ para un cierto $\lambda_{p+1} \in K$. Se deduce:

$$\varphi = \lambda_{p+1} \varphi_{p+1} + \sum_{i=1}^p \lambda_i \varphi_i \cdot \text{c.q.d.}$$

Si H es un subespacio de dimensión finita de E , se tiene también $H^{\circ\circ} = H$; lo que resulta, por ejemplo, de VIII.5.8 y de la inyectividad de la aplicación $E \rightarrow E^{**}$, asegurada por el teorema de Zorn.

Asimismo, VIII.5.7 es válido con E, F cualesquiera (si la $\dim [f(E)]$ es finita, la $\dim [{}^t f(E)]$ es también finita, y ambas dimensiones son iguales). La demostración descansa también en el axioma de la elección.

Nota sobre los módulos de tipo finito

Si se pretende generalizar los resultados de este capítulo (§§ 3, 4 y 5) a los *módulos de tipo finito* (e.d. que admiten un sistema de generadores finito), se obtienen las siguientes propiedades:

(VIII.3.1) no es verdadero. Por ejemplo, si $A = \mathbf{Z}/n\mathbf{Z}$ (en donde $n \in \mathbf{N}^*$ es cualquiera), el \mathbf{Z} -módulo A es de tipo finito, pues está engendrado por $\bar{1}$. Por lo tanto, no admite ninguna base, ya que todo elemento está ligado. Las consecuencias de VIII.3.1 no subsisten.

Por el contrario (VIII.3.3) subsiste en la forma siguiente, de la que más adelante daremos una demostración (cf. § X.1, p. 350): «si A es un anillo unífero conmutativo, y si M es un A -módulo que admite una base finita con n elementos, cualquier otra base de M tiene también n elementos». A un A -módulo M de esta clase se le llama *libre*, de *tipo finito* y de *dimensión* n . Vemos fácilmente que es isomorfo a A^n .

A pesar de todo, las propiedades que siguen a VIII.3.3 *no se conservan* para los módulos libres de tipo finito. El ejemplo sencillo que damos a continuación nos muestra las dificultades con las que nos tropezamos:

El \mathbf{Z} -módulo \mathbf{Z} es libre de dimensión 1, siendo $\{1\}$ una base. Todo elemento no nulo de \mathbf{Z} es \mathbf{Z} -libre, y el submódulo $2\mathbf{Z}$ es libre de dimensión 1, engendrado por $\{2\}$. No obstante la inclusión de $2\mathbf{Z}$ en \mathbf{Z} es estricta.

Si F es un submódulo libre de tipo finito del módulo de tipo finito E , F no admite necesariamente un suplementario. En el ejemplo que precede, $2\mathbf{Z}$ no admite suplementario alguno, puesto que la \mathbf{Z} -dimensión de \mathbf{Z} y de $2\mathbf{Z}$ es 1.

§ 4 El teorema VIII.4.2 subsiste (la misma demostración), pero VIII.4.3 no subsiste, y se generaliza en una forma algo más débil.

§ 5 Las definiciones VIII.5.1, 5.2 y 5.3 subsisten, y el teorema VIII.5.1 también (la misma demostración). Es posible definir la aplicación canónica J de un módulo en su bidual, pero en general no es inyectiva.

El teorema VIII.5.2 subsiste (para un módulo libre de tipo finito), así como el VIII.5.3. Los enunciados que siguen al VIII.5.3 no se pueden generalizar.

§ VIII.6 LENGUAJE DE LA GEOMETRÍA AFÍN

Suponemos que el cuerpo de base K es cualquiera.

DEFINICIÓN VIII.6.1

Sean T un espacio vectorial (por la izquierda) sobre K y E un conjunto. Definir en E una estructura de espacio afín ligada a T , consiste en dar una ley externa sobre E , de dominio T , designada por $(t, x) \mapsto t + x$, (o por $(t, x) \mapsto x + t$), tal que

(A₁) $\begin{cases} (t + t') + x = t + (t' + x) & \text{para } (t, t' \in T, x \in E), \\ 0 + x = x & \text{para } x \in E; \end{cases}$

(A₂) para todos $x, y \in E$, existe $t \in T$ tal que $y = t + x$;

(A₃) la relación $(t + x = 0 \text{ para todo } x \in E)$ implica $(t = 0)$.

(A₁) significa que el grupo aditivo de T opera sobre E (cf. Cap. II, § 8). Para todo $t \in T$, la aplicación $x \mapsto t + x$ es, pues, una biyección de E cuya biyección recíproca es $x \mapsto -t + x$. A esta biyección se le llama **traslación de vector t** : la designaremos por τ_t .

Designemos por \mathfrak{S}_E al grupo de las permutaciones de E . La aplicación $t \mapsto \tau_t$ es un homomorfismo del grupo aditivo de T en \mathfrak{S}_E ; (A₃) nos asegura que este homomorfismo es inyectivo, y permite identificar T aditivo con un grupo de biyecciones de E . Éste es el *grupo de las traslaciones* de E .

La aplicación $t \mapsto t + x$ de T en E , cuando $x \in E$ es fijo, es epiyectiva en virtud de (A₂), y veamos que es inyectiva. Para ello sea $t \in T$ tal que $t + x = x$; si $y \in E$, existe un $u \in T$ tal que $y = u + x$, de donde:

$$t + y = t + (u + x) = (t + u) + x = (u + t) + x = u + (t + x) = u + x = y,$$

es decir, $t + y = y$; así, se tiene $t + y = y$ para todo $y \in E$, de donde $t = 0$ según (A₃). Vemos, pues, que para todo $x \in E$, la aplicación $t \mapsto t + x$ de T en E es una biyección.

Estructuras de espacio vectorial definidas sobre un espacio afín

Sea (E, T) un espacio afín ligado al espacio vectorial T , y fijemos un elemento cualquiera $\omega \in E$. En virtud de lo que antecede, para todo $x \in E$, existe un $t \in T$, único, tal que $x = t + \omega$. Este elemento lo designaremos por $x - \omega$. La aplicación $x \mapsto x - \omega$ es una biyección de E en T . Cuando, por medio de esta biyección, se transporta la estructura de espacio vectorial de T a E , se obtiene sobre E una es-

estructura de espacio vectorial. Se dice que esta estructura *se ha obtenido tomando en E el origen ω* . La designaremos por medio de E_ω (en efecto, para esta estructura, ω es el elemento nulo de E). A veces resulta cómodo escribir $\overrightarrow{\omega x} = x - \omega$. Según (A_1) se tiene entonces la *fórmula de Chasles*:

$$\overrightarrow{a_1 a_2} + \overrightarrow{a_2 a_3} + \cdots + \overrightarrow{a_{n-1} a_n} = \overrightarrow{a_1 a_n},$$

cualesquiera que sean $a_1, a_2, \dots, a_n \in A$.

Aplicaciones afines. Grupo afín

DEFINICIÓN VIII.6.2

Sean (E, T) y (F, U) dos espacios afines, ligados respectivamente a los espacios vectoriales T y U . A una aplicación $\varphi: E \rightarrow F$ se le llama **afín** si existe un $f \in \mathcal{L}(T, U)$ y un $x \in E$ que verifiquen:

$$(1) \quad \varphi(t + x) = f(t) + \varphi(x) \text{ para todo } t \in T.$$

Se dice que la aplicación lineal f (que, como veremos, es única) está **asociada** a φ . Se dice también que f es la **parte lineal** de φ .

Si $y \in E$ es otro punto de E , existe un $u \in T$ tal que $y = x + u$. Se tiene:

$$\varphi(t + y) = \varphi(u + t + x) = f(u + t) + \varphi(x), \text{ según (1),}$$

o sea:

$$\varphi(t + y) = f(u) + f(t) + \varphi(x);$$

pero, siempre en virtud de (1),

$$\varphi(y) = \varphi(u + x) = f(u) + \varphi(x), \text{ de donde } \varphi(t + y) = f(t) + \varphi(y).$$

En otras palabras, si (1) se verifica para un punto $x \in E$, se verifica para todos los puntos $x \in E$. Esto demuestra que la aplicación f es única.]]

VIII.6.1 La compuesta de dos aplicaciones afines es una aplicación afín.

Demostración. Sean (E, T) , (F, U) , (G, V) tres espacios afines,

$$\varphi: E \rightarrow F \text{ y } \psi: F \rightarrow G$$

aplicaciones afines, cuyas aplicaciones lineales asociadas son, respectivamente,

$$f \in \mathcal{L}(T, U) \quad \text{y} \quad g \in \mathcal{L}(U, V)$$

Se tiene:

$$\psi \circ \varphi(t + x) = \psi[f(t) + \varphi(x)] = g[f(t) + \psi[\varphi(x)]] = g \circ f(t) + \psi \circ \varphi(x).$$

Luego $\psi \circ \varphi$ es afín, y su aplicación lineal asociada es $g \circ f$. c.q.d.

VIII.6.2 *Para que la aplicación afín $\varphi : E \rightarrow F$ definida por (1) sea biyectiva, es necesario y suficiente que f sea biyectiva.*

Demostración. En (1), se fija x , y la proposición resulta del hecho de que $t \rightarrow t + x$ y $u \rightarrow u + \varphi(x)$ sean, respectivamente, biyecciones de T y U en E y F .]

La aplicación idéntica del espacio afín (E, T) es afín. De todo lo anterior, se deduce:

VIII.6.3 *El conjunto de las biyecciones afines de un espacio afín (E, T) en sí mismo, forma un grupo de biyecciones de E , llamado el **grupo afín de E** y se designa por $\mathcal{A}(E)$. La aplicación $\mathcal{A}(E) \rightarrow \text{GL}(T)$ que, a cada $\varphi \in \mathcal{A}(E)$ asocia su **parte lineal** f es un homomorfismo de grupos. El núcleo de este homomorfismo es el **grupo de las traslaciones de E** .*

La *Geometría afín* es el estudio de las propiedades de las figuras de un espacio afín, que son invariantes en su grupo afín.

Ejemplo fundamental de espacio afín

Se considera un espacio vectorial T , se toma $E = T$. La acción de T sobre T se define por la ley de grupo de T : $(t, x) \mapsto t + x$. Las condiciones (A_1) , (A_2) y (A_3) se verifican trivialmente. El espacio afín así obtenido se designará con el nombre de «espacio afín T ». Es obvio que no se debe confundir con el espacio vectorial T .

Una aplicación afín $\varphi : T \mapsto T$ es simplemente una aplicación de la forma:

$$\varphi(x) = b + f(x) \quad (\text{haciendo en (1), } x = t \text{ y } b = \varphi(0)), \text{ en donde } f \in \mathcal{L}(T).$$

El *grupo afín* es el subgrupo de las biyecciones de T engendrado por el grupo $\text{GL}(T)$ y por el grupo aditivo T de las traslaciones de T .

Si (E, T) designa un espacio afín ligado a T , y ω un punto cualquiera de E , se puede, en particular, dotar al espacio vectorial E_ω de la estructura definida ante-

riormente. La biyección canónica $E_\omega \rightarrow E'$ es una biyección afín. Luego el espacio afín (E, T) es isomorfo al espacio afín (E_ω, T) , es decir, a T . (Sin embargo, ninguno de los isomorfismos así definidos es privilegiado, por lo que se dice que no existe isomorfismo canónico entre el espacio afín E y el espacio afín T .)

Para demostrar las propiedades de un espacio afín E ligado a T , bastará con demostrarlas para el propio espacio afín T .

Subvariedades afines

DEFINICIÓN VIII.6.3

Una subvariedad afín V de un espacio afín (E, T) es el conjunto vacío, o bien un subespacio vectorial de un espacio vectorial E_a , en donde $a \in E$.

Si V es una variedad afín no vacía de E , y es un subespacio vectorial de E_a , se tiene $a \in V$. Sea entonces $b \in V$ un elemento cualquiera, la traslación \overrightarrow{ab} hace corresponder b a a , y es un isomorfismo de E_a en E_b ⁽¹⁾. Luego V es un subespacio vectorial de E_b , para todo $b \in V$.

Además, después de lo que precede, todos los subespacios vectoriales de los $(E_b)_{b \in V}$, iguales a V , son isomorfos dos a dos. A su dimensión común se le llama *dimensión de V* , y se designa por $\dim(V)$. Por definición, $\dim(\emptyset) = -1$.

A su codimensión común se le llama *codimensión de V* , se designa por la $\text{codim}(V)$

$\dim(V) = 0$ si, y sólo si, V se reduce a un punto de E .

Si $\dim(V) = 1$, se dice que V es una *recta afín*.

Si $\dim(V) = 2$, se dice que V es un *plano afín*.

Si T es de dimensión finita n , y si $\dim(V) = n - 1$, se dice que V es un *hiperplano afín*. Si $\dim(V_1)$ y $\dim(V_2)$ son finitos e iguales, y si $V_1 \subset V_2$, se tiene $V_1 = V_2$ (cf. VIII.3.7).

En general, si T es cualquiera, se dice que V es un *hiperplano afín* si $\text{codim}(V) = 1$.

VIII.6.4 La intersección de una familia cualquiera de subvariedades afines del espacio afín (E, T) es una subvariedad afín.

Demostración. Sea $(V_i)_{i \in I}$ una familia de subvariedades afines. Si $\bigcap_{i \in I} V_i = \emptyset$, la proposición está demostrada. Si no, sea $a \in \bigcap_{i \in I} V_i$. Cada V_i es un subespacio vectorial de E_a , por lo tanto también lo es $\bigcap_{i \in I} V_i$. c.q.d.

⁽¹⁾ Este isomorfismo expresa la «regla del paralelogramo». Cf. tomo 3, *Geometría*.

VIII.6.4 permite definir la noción de *variedad afín engendrada*:

DEFINICIÓN VIII.6.4

$\left\{ \begin{array}{l} A \text{ la intersección de subvariedades afines de } E, \text{ que contienen a una parte } A \\ \text{de } E, \text{ se le llama variedad afín engendrada por } A \text{ y se designa por } \text{Af}(A). \end{array} \right.$

— La imagen directa (resp. recíproca) de un subespacio afín por una aplicación afín, es un subespacio afín.

Dirección de una subvariedad afín

Sea V una subvariedad afín de E y sea $a \in V$; V es un subespacio vectorial de E_a ; la biyección natural, definida antes, de E_a en T , transforma V en un subespacio vectorial V_0 de T . Se ve fácilmente que V_0 no depende del punto a , y por definición, V_0 es la *dirección* de V .

A dos variedades afines, V, W se les llama **paralelas** si $V_0 \subset W_0$ o si $W_0 \subset V_0$. Cuando esto ocurre, se tiene $V \subset W$, o bien $W \subset V$, o $V \cap W = \emptyset$.

Subvariedades afines del espacio afín T (T designa un K -espacio vectorial)

Por definición, una variedad afín V no vacía de T es la transformada por una traslación $t \in T$ de un subespacio vectorial V_0 de T . V_0 es precisamente la *dirección* de V . Escribiremos $V = t + V_0$. Supongamos que T es de dimensión finita. Si $f_i(x) = 0$ ($1 \leq i \leq r$) designa un sistema de ecuaciones de V_0 , V se halla definido por el sistema de ecuaciones $f_i(x) = b_i$ ($= f_i(a)$), en donde $a \in V$ es fijo ($1 \leq i \leq r$).

En particular, un *hiperplano afín* de T se halla definido por una ecuación de la forma $f(x) = b$, en donde f designa una forma lineal sobre T y b es un escalar.

Sistema afín libre

Sea Λ una parte de un espacio afín (E, T) . Para todo par de puntos (a, b) de Λ , vemos que $\Lambda_a = \Lambda \setminus \{a\}$ es una parte libre de E_a , si, y sólo si, $\Lambda_b = \Lambda \setminus \{b\}$ es una parte libre de E_b . De la relación $\sum_{x \in \Lambda, x \neq a} \lambda_x \cdot \overrightarrow{ax} = 0$, se deduce $\sum_{x \in \Lambda, x \neq a} \lambda_x (\overrightarrow{ab} + \overrightarrow{bx}) = 0$, de donde

$$\sum_{y \in \Lambda, y \neq b} \mu_y \cdot \overrightarrow{by} = 0, \quad \text{con} \quad \mu_y = \lambda_y \quad \text{si} \quad y \neq a, \quad \text{y} \quad \mu_a = - \sum_{x \neq a} \lambda_x.$$

Si los λ_x no son todos nulos, los (μ_y) tampoco serán todos nulos de donde resulta nuestra proposición. Esta propiedad nos conduce a la siguiente:

DEFINICIÓN VIII.6.5

*Una parte A del espacio afín (E, T) es **libre desde el punto de vista afín** si el conjunto $\Lambda \setminus \{a\}$ es una parte libre de E_a , en donde el origen a es uno de los puntos de Λ (en tal caso lo mismo ocurre para **todo** punto $a \in \Lambda$). A una parte afín libre Λ de E tal que $\text{Af}(\Lambda) = E$ se le llama **base afín** de E .*

Consecuencias

Toda parte de una parte afín libre es afín libre. Si T es de dimensión finita n , toda parte afín libre de E tiene a lo sumo $n + 1$ elementos; y una parte afín libre de E es una base afín si, y sólo si, tiene exactamente $n + 1$ elementos. La variedad afín engendrada por una parte afín libre de $p + 1$ elementos es de dimensión p .

Coordenadas baricéntricas: ver tomo 3 (Geometría).

Capítulo IX

Matrices

- En todo este capítulo, se supondrá *conmutativo* el cuerpo de base K . Se recuerda que \mathbf{N}_n^* designa al conjunto de los n primeros enteros > 0 .

§ IX.1 MATRICES

DEFINICIÓN IX.1.1

Se llama **matriz de tipo (n, p)** (o **(n, p) -matriz** con coeficientes en K , a toda aplicación de $\mathbf{N}_n^* \times \mathbf{N}_p^n$ en K .

Una (n, p) -matriz es, pues, una aplicación que asigna a cada entero $i \leq n$ y a cada entero $j \leq p$, un elemento a_{ij} de K . Al a_{ij} se le llama *término general* de M , y a los a_{ij} también se les llama *coeficientes* de M .

Notaciones

$$M = [a_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \quad \text{o} \quad M = [a_{ij}]$$

$$M =_i \begin{bmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,p} \\ \vdots & & \vdots & & \vdots \\ \dots & & a_{ij} & & \dots \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \dots & \dots & \dots & a_{n,p} \end{bmatrix}.$$

La fila de índice i de M es la sucesión $(a_{i,1}, \dots, a_{i,p})$.

La columna de índice j de M es la sucesión $(a_{1,j}, \dots, a_{n,j})$.

M contiene n filas y p columnas. El único elemento común a la fila i y a la columna j es a_{ij} . Al índice i se le llama *índice-fila*, y al índice j se le llama *índice-columna*. Cada fila [resp. cada columna] de M se puede identificar con un vector de K^n llamado *vector fila* [resp. un vector de K^p llamado *vector columna*] de M .

El conjunto de las (n, p) -matrices con elementos en K se designa por $\mathcal{M}_{n,p}(K)$. Las (n, n) -matrices son las llamadas *matrices cuadradas de orden n* . Se escribe $\mathcal{M}_n(K)$ o $M_n(K)$, en vez de $\mathcal{M}_{n,n}(K)$.

En una matriz cuadrada $[a_{ij}]$, a los a_{ij} cuyos índices son iguales se les llama términos de la *diagonal principal*.

Una matriz **diagonal** es una matriz cuadrada $[sa_{ij}]$ tal que $a_{ij} = 0$ para $i \neq j$.

En $\mathcal{M}_{n,p}(K)$ definimos una *adición* de la manera siguiente:

$$(1) \quad \begin{aligned} M &= [a_{ij}] \text{ y } N = [b_{ij}] \\ M + N &= [a_{ij} + b_{ij}] . \end{aligned}$$

Definimos también una ley externa, de dominio K , estableciendo:

$$(2) \quad \begin{aligned} M &= [a_{ij}] , \text{ y } \lambda \in K , \\ \lambda M &= [\lambda a_{ij}] . \end{aligned}$$

El teorema que sigue es entonces evidente:

TEOREMA IX.1.1

|| El conjunto $\mathcal{M}_{n,p}(K)$, dotado de las leyes definidas por las fórmulas (1) y (2), es un K -espacio vectorial.

El elemento nulo de $\mathcal{M}_{n,p}(K)$ es la matriz que tiene nulos todos sus términos. Dicha matriz se designa por 0. La matriz *opuesta* de $M = [a_{ij}]$ es $-M = [-a_{ij}]$.

Nota. Recordemos que si E es un conjunto cualquiera, el *símbolo de Kronecker de orden k* con valores en K es la aplicación

$$\delta : E^k \rightarrow K, \quad (u_1 \dots u_k) \mapsto \delta_{u_1, u_2, \dots, u_k}$$

tal que $\delta_{u_1, u_2, \dots, u_k} = 1$ si $u_1 = u_2 = \dots = u_k$, y $\delta_{u_1, \dots, u_k} = 0$ si existe un i y un j tales que $u_i \neq u_j$. Cuando $k = 2$ y $E = \mathbf{N}_n^*$, se obtiene el símbolo de Kronecker

habitual $\delta_{i,j}$. Cuando $k = 2$ y $E = \mathbf{N}_n^* \times \mathbf{N}_p^*$, el símbolo de Kronecker se designa por

$$\delta_{ij,kl} \quad ((i,j) \in E, (k,l) \in E).$$

Matrices elementales

Por definición, la matriz elemental E_{ij} es la matriz

$$E_{ij} = [\delta_{ij,kl}]_{\substack{1 \leq k \leq n \\ 1 \leq l \leq p}};$$

en otras palabras, el elemento a_{ij} de E_{ij} vale 1, y los restantes elementos son nulos.

La (n, p) -matriz $M = [a_{ij}]$ admite una expresión única de la forma

$$(3) \quad M = \sum_{1 \leq i \leq n, 1 \leq j \leq p} a_{ij} E_{ij}.$$

Podemos, pues, enunciar:

IX.1.2 *El espacio vectorial $\mathcal{M}_{n,p}(K)$ es de dimensión np y admite por base el conjunto de las matrices elementales E_{ij} ($1 \leq i \leq n$, $1 \leq j \leq p$).*

Trasposición

La *traspuesta* de una (n, p) -matriz $M = [a_{ij}]$ es la matriz, que designaremos por tM , definida por

$${}^tM = [b_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}, \quad b_{ij} = a_{ji} \text{ para todo } i \text{ y todo } j.$$

tM es, pues, una (p, n) -matriz. Se tiene: ${}^t({}^tM) = M$.

Es inmediato que la aplicación $M \mapsto {}^tM$ es un isomorfismo del espacio vectorial $\mathcal{M}_{n,p}(K)$ en el espacio vectorial $\mathcal{M}_{p,n}(K)$.

Producto de matrices

DEFINICIÓN IX.1.2

El **producto** de la (n, p) -matriz $M = [a_{ij}]$ y de la (p, q) -matriz $N = [b_{kl}]$, es la (n, q) -matriz, que se designa por $M.N$ o MN , y está definida por

$$MN = [c_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}} \quad c_{ij} = \sum_{k=1}^p a_{ik} b_{kj} \text{ para } i \in \mathbf{N}_n^* \text{ y } j \in \mathbf{N}_q^*.$$

Propiedades del producto

1) Si $M \in \mathcal{M}_{n,p}(K)$, la aplicación $N \mapsto M \cdot N$ de $\mathcal{M}_{p,q}(K)$ en $\mathcal{M}_{n,q}(K)$ es lineal. Análogamente, para toda $N \in \mathcal{M}_{p,q}(K)$, la aplicación $M \mapsto M \cdot N$ de $\mathcal{M}_{n,p}(K)$ en $\mathcal{M}_{n,q}(K)$ es lineal. Dicho de otra manera, se verifican las fórmulas de distributividad del producto respecto de la adición:

$$M(N_1 + N_2) = MN_1 + MN_2, \quad (M_1 + M_2)N = M_1N + M_2N,$$

y la fórmula

$$\lambda(MN) = (\lambda M)N = M(\lambda N)$$

$$(M, M_1, M_2 \in \mathcal{M}_{n,p}(K), \quad N, N_1, N_2 \in \mathcal{M}_{p,q}(K), \quad \lambda \in K).$$

2) Asociatividad

Sean $M = [a_{ij}]$, $N = [b_{kl}]$, $P = [c_{mn}]$, respectivamente, una (n, p) -matriz, una (p, q) -matriz y una (q, r) -matriz.

Entonces los productos $M(NP)$ y $(MN)P$ están definidos, y se verifica:

$$M(NP) = (MN)P.$$

3) Cuando el producto MN está definido, ${}^tN {}^tM$ también lo está, y se verifica:

$${}^t(MN) = {}^tN {}^tM.$$

Obsérvese que el producto MN solamente está definido cuando el número de columnas de M coincide con el número de filas de N .

Si M es una (n, p) -matriz y N es una (p, n) -matriz, los productos MN y NM están ambos definidos. En este caso, las matrices MN y NM son ambas cuadradas, de órdenes respectivos n y p .

Comprobación de la propiedad 2)

$$(MN) = [d_{ij}], \quad \text{con} \quad d_{ij} = \sum_{k=1}^p a_{ik} b_{kj};$$

$$(MN) \cdot P = [f_{ij}], \quad \text{con} \quad f_{ij} = \sum_{k=1}^q d_{ik} c_{kj}; \quad d_{ik} = \sum_{l=1}^p a_{il} b_{lk};$$

de donde:

$$f_{ij} = \sum_{\substack{1 \leq k \leq q \\ 1 \leq l \leq p}} a_{il} b_{lk} c_{kj} = \sum_{\substack{1 \leq k \leq p \\ 1 \leq l \leq q}} a_{ik} b_{kl} c_{lj}.$$

Análogamente se demuestra que el término general de $M(NP)$ viene dado por la fórmula anterior. c.q.d.

Comprobación de la propiedad 3)

Si $M = [a_{ij}]$, $N = [b_{ij}]$, se tiene:

$$M \cdot N = [c_{ij}], \quad \text{con} \quad c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}.$$

Pongamos ${}^tM = [\alpha_{ij}]$, ${}^tN = [\beta_{ij}]$ y ${}^tN \cdot {}^tM = [\gamma_{ij}]$. Se tiene:

$$\gamma_{ij} = \sum_{k=1}^p \beta_{ik} \alpha_{kj}.$$

Además: $\alpha_{kj} = a_{jk}$, $\beta_{ik} = b_{ki}$, de donde:

$$\gamma_{ij} = \sum_{k=1}^p b_{ki} a_{jk},$$

y puesto que hemos supuesto que K es conmutativo,

$$\gamma_{ij} = \sum_{k=1}^p a_{jk} b_{ki} = c_{ji}, \quad \text{de donde} \quad {}^tN \cdot {}^tM = {}^t(MN) \quad \text{c.q.d.}$$

Estructura algebraica de $\mathcal{M}_n(K)$

El producto de matrices es una ley interna en $\mathcal{M}_n(K)$, que admite a la (n, n) -matriz $I_n = [\delta_{i,j}]$ ($\delta_{i,j}$: símbolo de Kronecker)

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix}.$$

como elemento neutro.

De las propiedades 1) y 2) establecidas anteriormente se deduce el teorema siguiente:

TEOREMA IX.1.3

Si dotamos al K -espacio vectorial $\mathcal{M}_n(K)$ de la ley interna

$$(M, N) \mapsto M \cdot N \text{ (producto de matrices),}$$

$\mathcal{M}_n(K)$ es una K -álgebra, cuyo elemento unidad es I_n .

El homomorfismo $K \rightarrow \mathcal{M}_n(K) : \lambda \mapsto \lambda I_n$ es inyectivo. Su imagen es un subanillo de $\mathcal{M}_n(K)$, isomorfo a K , formado por las matrices de la forma

$$\begin{bmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & & \\ \vdots & & \ddots & \\ 0 & & & \lambda \end{bmatrix}.$$

A estas matrices se les llama *matrices escalares*.

- Si $n = 1$, $\mathcal{M}_n(K)$ se puede identificar con K .
- Si $n > 1$, el anillo $\mathcal{M}_n(K)$ jamás es íntegro. Por ejemplo, si

$$M = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{y} \quad N = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \text{ se tiene } M \cdot N = 0.$$

— $\mathcal{M}_n(K)$ jamás es conmutativo para $n > 1$. Por ejemplo, si M y N son las anteriores matrices, se tiene:

$$N \cdot M = P, \quad \text{con} \quad P = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Los elementos invertibles del anillo $\mathcal{M}_n(K)$ se llaman *matrices invertibles de orden n* (con elementos en A). Su conjunto constituye el grupo de los elementos invertibles de $\mathcal{M}_n(K)$:

DEFINICIÓN IX.1.3

$\}$ El grupo de los *elementos invertibles* de $\mathcal{M}_n(K)$ se llama *grupo lineal de orden n sobre K* , y se designa por $GL_n(K)$ o $GL(n, K)$.

A los grupos $GL_n(K)$, y sus subgrupos notables, se les llama *grupos clásicos*. Su estudio, inaugurado por Jordan (1838-1922), constituye una parte importante del Álgebra.

Suspensión de $\mathcal{M}_n(K)$ en $\mathcal{M}_p(K)$ ($p \geq n$)

Sea $\{E_{ij}\}_{1 \leq i, j \leq n}$ la base canónica de $\mathcal{M}_n(K)$. A toda matriz $M = [a_{ij}]$ cuadrada de orden n , le asociamos la matriz $j_{n,p}(M)$, cuadrada de orden p , tal que:

$$j_{n,p}(M) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} E_{ij}.$$

En otras palabras, $j_{n,p}(M)$ es la matriz:

$$\left[\begin{array}{ccc|ccc} \overbrace{\begin{matrix} M \end{matrix}}^n & 0 & \dots & 0 & & \\ & \vdots & & \vdots & & \\ & 0 & \dots & 0 & & \\ & \vdots & & \vdots & & \\ & 0 & \dots & \dots & \ddots & \\ & 0 & \dots & \dots & 0 & \end{array} \right]_p.$$

La aplicación $M \mapsto j_{n,p}(M)$ es lineal, inyectiva, y tal que

$$j_{n,p}(M \cdot N) = j_{n,p}(M) \cdot j_{n,p}(N),$$

pero no transforma I_n en I_p . Su imagen es un anillo unífero para las leyes de $\mathcal{M}_p(K)$, pero su elemento unidad $j_{n,p}(I_n)$ es distinto de I_p . No es, pues, un subanillo unífero de $\mathcal{M}_p(K)$. A $j_{n,p}$ se le llama *aplicación de suspensión*, y evidentemente, para $n < p < q$, $j_{p,q} \circ j_{n,p} = j_{n,q}$. Observemos, finalmente, que toda matriz de la forma $j_{n,p}(M)$ ($p > n$) es un divisor de cero en $\mathcal{M}_p(K)$.

Suspensión de $GL_n(K)$ en $GL_p(K)$ ($p > n$)

Sea nuevamente $\{E_{ij}\}$ la base canónica de $\mathcal{M}_p(K)$. A toda matriz $M = [a_{ij}]$ cuadrada de orden n , le asociamos la matriz $\psi_{n,p}(M)$ cuadrada de orden p tal que:

$$\psi_{n,p}(M) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} E_{ij} + \sum_{k=n+1}^p E_{k,k}.$$

En otras palabras, $\psi_{n,p}(M)$ es la matriz:

$$\left[\begin{array}{ccc|ccc} \overbrace{\begin{matrix} M \end{matrix}}^n & 0 & \dots & 0 & & \\ & \vdots & & \vdots & & \\ & 0 & & \vdots & & \\ & 0 & & \vdots & & \\ & 0 & \dots & 1 & & \\ & \vdots & & \vdots & & \\ & 0 & \dots & \dots & \ddots & \\ & 0 & \dots & \dots & 0 & 1 \end{array} \right].$$

La aplicación $\psi_{n,p}$ no es lineal, pero verifica:

$$\psi_{n,p}(M \cdot N) = \psi_{n,p}(M) \psi_{n,p}(N), \quad \text{y} \quad \psi_{n,p}(I_n) = I_p;$$

además $\psi_{n,p}$ es inyectiva.

En particular, la restricción $\varphi_{n,p}$ de $\psi_{n,p}$ a $GL_n(K)$ es un homomorfismo inyectivo de $GL_n(K)$ en $GL_p(K)$ (llamado «de suspensión»).

Con la ayuda de $\varphi_{n,p}$, se puede identificar $GL_n(K)$ con un subgrupo de $GL_p(K)$. Es evidente que las suspensiones se componen:

$$\text{para } n \leq p \leq q, \quad \varphi_{n,q} = \varphi_{p,q} \circ \varphi_{n,p}.$$

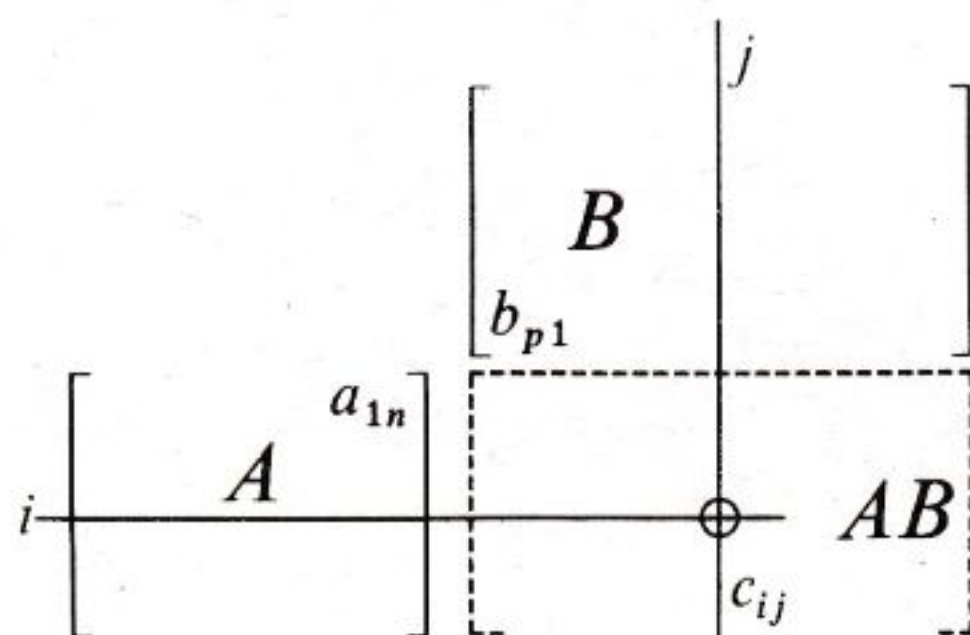
Ejemplos de cálculo con matrices

Indiquemos ante todo la *disposición práctica* del producto de matrices, señalado por M. Martin; para obtener el producto AB , en donde

$$A = [a_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \quad \text{y} \quad B = [b_{jk}]_{\substack{1 \leq j \leq p \\ 1 \leq k \leq q}},$$

se aproximan los términos a_{1n} y b_{p1} (ver esquema).

El término c_{ij} de AB se obtiene en la intersección de la fila i de A y de la columna j de B , y es el «producto escalar» del vector fila de índice i de A y del vector columna de índice j de B .



El procedimiento es particularmente cómodo cuando se debe efectuar el producto de, por lo menos, tres matrices A, B, C, \dots ; se dispone (AB) y C de la misma forma que A y B (sin tener necesidad de volver a escribir AB); como ejemplo, calculemos ABC , en donde

$$A = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix}$$

$$y \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\operatorname{sen} \psi \\ 0 & \operatorname{sen} \psi & \cos \psi \end{bmatrix}, \quad (\text{matrices asociadas a rotaciones de } \mathbf{R}^3)$$

$$\begin{array}{c} A \\ \left[\begin{array}{ccc} \cos \theta & -\operatorname{sen} \theta & 0 \\ \operatorname{sen} \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{array} \right] \end{array} \begin{array}{c} B \\ \left[\begin{array}{ccc} \cos \varphi & 0 & -\operatorname{sen} \varphi \\ 0 & 1 & 0 \\ \operatorname{sen} \varphi & 0 & \cos \varphi \end{array} \right] \end{array} \begin{array}{c} C \\ \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \cos \psi & -\operatorname{sen} \psi \\ 0 & \operatorname{sen} \psi & \cos \psi \end{array} \right] \end{array} \begin{array}{c} ABC \\ \left[\begin{array}{ccc} \cos \theta \cos \varphi, -\operatorname{sen} \theta \cos \psi - \cos \theta \operatorname{sen} \varphi \operatorname{sen} \psi, \operatorname{sen} \theta \operatorname{sen} \psi - \cos \theta \cos \psi \operatorname{sen} \varphi \\ \operatorname{sen} \theta \cos \varphi, \cos \theta \cos \psi - \operatorname{sen} \theta \operatorname{sen} \varphi \operatorname{sen} \psi, -\cos \theta \operatorname{sen} \psi - \operatorname{sen} \theta \operatorname{sen} \varphi \cos \psi \\ \operatorname{sen} \varphi, \cos \varphi \operatorname{sen} \psi, \cos \varphi \cos \psi \end{array} \right] \end{array}$$

1) En el anillo $\mathcal{M}_n(K)$, busquemos los productos de las matrices elementales. Se tiene:

$$E_{ij} = [\delta_{ij,rs}]_{\substack{1 \leq r \leq n \\ 1 \leq s \leq n}}; \quad E_{kl} = [\delta_{kl,uv}]_{\substack{1 \leq u \leq n \\ 1 \leq v \leq n}},$$

de donde

$$E_{ij} E_{kl} = [C_{\lambda\mu}]_{\substack{1 \leq \lambda \leq n \\ 1 \leq \mu \leq n}} \quad \text{con} \quad C_{\lambda\mu} = \sum_{s=1}^n \delta_{ij,\lambda s} \delta_{kl,s\mu}.$$

En la expresión de $C_{\lambda\mu}$, el único término no nulo se obtiene cuando

$$i = \lambda, \quad j = s = k, \quad l = \mu.$$

$$\text{si } j \neq k: E_{ij} E_{kl} = 0 \quad \text{y} \quad (\text{si } j = k): E_{ik} E_{kl} = E_{il},$$

que se puede escribir de la manera siguiente:

$$(1) \quad E_{ij} E_{kl} = \delta_{jk} E_{il}.$$

2) Matrices triangulares

Una matriz cuadrada $M \in \mathcal{M}_n(K)$ se llama *triangular superior* si (escribiendo $M = [a_{ij}]$) se tiene $a_{ij} = 0$ para $i < j$, y es *triangular inferior* si $a_{ij} = 0$ para $i > j$.

Es claro que el conjunto de las matrices triangulares superiores (resp. inferiores) es un subespacio vectorial de $\mathcal{M}_n(K)$, de dimensión $\frac{n(n+1)}{2}$, puesto que es el subespacio engendrado por las matrices elementales E_{ij} , $i \geq j$ (resp. $i \leq j$).

En el ejemplo 1), hemos visto que $E_{ij} E_{i'j'} = \delta_{ji'} E_{ij'}$.

Sabemos que $N^4 = 0$. Además

$$N^2 = \begin{bmatrix} 0 & 0 & ad & ae + bf \\ 0 & 0 & 0 & df \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad N^3 = \begin{bmatrix} 0 & 0 & 0 & adf \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Por otro lado, puesto que I_4 y N conmutan, se puede (para $m \geq 4$) aplicar la fórmula del binomio que en este caso se reduce a:

$$(I_4 + N)^m = I_4 + mN + \frac{m(m-1)}{2} N^2 + \frac{m(m-1)(m-2)}{6} N^3$$

(puesto que $N^k = 0$ para $k \geq 4$). Se deduce:

$$M^m = \begin{bmatrix} 1 & ma & mb + \frac{m(m-1)}{2} ad & mc + \frac{m(m-1)}{2} (ae + bf) + \frac{m(m-1)(m-2)}{6} adf \\ 0 & 1 & md & me + \frac{m(m-1)}{2} df \\ 0 & 0 & 1 & mf \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

4) Una matriz cuadrada $M = [a_{ij}]$ se llama *simétrica* si ${}^tM = M$; *antisimétrica* si ${}^tM = -M$. El conjunto de las matrices simétricas (resp. antisimétricas) forma un subespacio vectorial de $\mathcal{M}_n(K)$, que designaremos por $S_n(K)$ (resp. $A_n(K)$).

Se tiene:

$$2M = (M + {}^tM) + (M - {}^tM).$$

Es evidente que $M + {}^tM$ es simétrica y que $M - {}^tM$ es antisimétrica. Si el cuerpo K es de característica $\neq 2$, se deduce:

$$M = \frac{1}{2}(M + {}^tM) + \frac{1}{2}(M - {}^tM), \text{ de donde } \mathcal{M}_n(K) = S_n(K) + A_n(K).$$

Puesto que en este caso, $(2u = 0) \Rightarrow (u = 0)$ (en donde $u \in K$), vemos además que $S_n(K) \cap A_n(K) = \{0\}$. Dicho de otra forma, si K es de característica $\neq 2$, $\mathcal{M}_n(K)$ es suma directa de $S_n(K)$ y de $A_n(K)$. Si $M = [a_{ij}]$ es simétrica, se tiene la descomposición:

$$M = \sum_{i < j} a_{ij}(E_{ij} + E_{ji}) + \sum_i a_{ii} E_{ii}.$$

Luego, las $\frac{n(n+1)}{2}$ matrices $(E_{ij} + E_{ji})_{i < j}$ y $(E_{ii})_{1 \leq i \leq n}$ forman una base de $S_n(A)$.

Asimismo, vemos que (cuando K es de característica $\neq 2$) las $\frac{n(n-1)}{2}$ matrices $(E_{ij} - E_{ji})_{i < j}$ forman una base de $A_n(K)$.

(I, J) -matrices

Sean I y J dos conjuntos finitos, de cardinal ≥ 1 . Se llama (I, J) -matriz con coeficientes en K a toda aplicación $I \times J \rightarrow K, (i, j) \mapsto K$. A una (I, I) -matriz se le llama simplemente I -matriz cuadrada.

Con estas definiciones, es posible entender la mayor parte de las definiciones y resultados de este capítulo. Puesto que esta extensión funciona, en adelante nos limitaremos a estudiar las (n, p) -matrices (caso en que $I = \mathbf{N}_n^*$ y $J = \mathbf{N}_p^*$).

Señalemos que la consideración de las (I, J) -matrices, con conjuntos I y J no identificables de una manera natural a los \mathbf{N}_k^* , es indispensable en álgebra multilineal.

§ IX.2 MATRICES Y APLICACIONES LINEALES

En este §, el cuerpo de base K se supone, siempre conmutativo.

DEFINICIÓN IX.2.1

Sean E un K -espacio vectorial de dimensión finita n , (e_1, \dots, e_n) una base finita (salida), F un K -espacio vectorial de dimensión finita y (f_1, \dots, f_p) una base (llegada); y sea $\varphi: E \rightarrow F$ una aplicación lineal. La matriz asociada a φ en las bases (e_i) y (f_j) es la (p, n) -matriz $M(\varphi) = [a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$, con elementos en K tal que

$$\varphi(e_j) = \sum_{i=1}^p a_{ij} f_i \quad (1 \leq j \leq n).$$

Cuando $E = F$ y se elige $f_i = e_i$ para todo i , a $M(\varphi)$ se le llama matriz del endomorfismo φ en la base (e_i) . La designaremos también por $M\varphi$.

● La matriz $M(\varphi)$ se obtiene, pues, escribiendo *en columna* las componentes de los $\varphi(e_j)$ en la base (f_1, f_2, \dots, f_p) :

$$\begin{array}{c} \varphi(e_1) \quad \dots \quad \varphi(e_j) \quad \dots \quad \varphi(e_n) \\ \begin{array}{c} f_1 \\ \vdots \\ f_i \\ \vdots \\ f_p \end{array} \left[\begin{array}{cccc} a_{11} & \dots & \cdot & a_{1n} \\ a_{21} & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ a_{p1} & \dots & \dots & a_{pn} \end{array} \right] \end{array} .$$

De las definiciones resulta que, si $(x_i)_{1 \leq i \leq n}$ son las coordenadas de $x \in E$ en la base (e_i) , y si $(y_i)_{1 \leq i \leq p}$ son las coordenadas de $y = \varphi(x)$ en la base (f_j) de F , se tienen las siguientes fórmulas de transformación:

(T)

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad 1 \leq i \leq p .$$

Es evidente que $\varphi \mapsto M(\varphi)$ es una aplicación lineal inyectiva de $\mathcal{L}_K(E, F)$ en $\mathcal{M}_{p,n}(K)$; en otras palabras, se tiene:

$$\begin{aligned} M(\varphi_1 + \varphi_2) &= M(\varphi_1) + M(\varphi_2) , \\ M(\lambda\varphi) &= \lambda M(\varphi) . \end{aligned}$$

Recíprocamente, sea $M = [a_{ij}]$ una (p, n) -matriz. Definamos una aplicación

$\varphi : E \rightarrow F$, por $\varphi(e_j) = \sum_{i=1}^p a_{ij} f_i$ ($1 \leq j \leq n$), y por

$$(1) \quad \varphi(x) = \varphi\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j \varphi(e_j) \quad \text{para todo elemento } x \in E$$

(se dice que φ se ha obtenido a partir de los $\varphi(e_j)$ «prolongando por linealidad»). φ está bien definida (ya que (e_j) es una base); φ es lineal (pues $x \mapsto (x_j)_{1 \leq j \leq n}$ es un isomorfismo de E en K^n), y se tiene por definición de φ : $M(\varphi) = M$.

En resumen, hemos demostrado el siguiente:

TEOREMA IX.2.1

|| Sean E y F dos K -espacios vectoriales de bases finitas $(e_i)_{1 \leq i \leq n}$ y $(f_j)_{1 \leq j \leq p}$.
|| La aplicación de $\mathcal{L}_K(E, F)$ en $\mathcal{M}_{p,n}(K)$ que, a cada aplicación lineal

|| *hace corresponder su matriz asociada en las bases (e_i) y (f_j) , es un isomorfismo de K -espacios vectoriales.*

Expresión matricial de una transformación lineal

Supongamos que se verifican las hipótesis del teorema IX.2.1. Con la ayuda de la base (e_i) (resp. (f_j)) identificamos E con K^n (resp. F con K^p). Un elemento (x_1, \dots, x_m) de K^m se representará por medio de una matriz columna:

$$(2) \quad \mathcal{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Sea $M_\varphi = [a_{ij}]$ la matriz asociada (en las bases anteriores) a la aplicación lineal $\varphi : E \rightarrow F$.

Las relaciones (1) se expresan, si $x \in E$ está representado por medio de la matriz (2):

$$\begin{aligned} \varphi(x) &= \sum_{j=1}^n x_j \sum_{i=1}^p a_{ij} f_i = \sum_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} a_{ij} x_j f_i \quad (\text{pues } K \text{ es conmutativo}) \\ &= \sum_{i=1}^p \left(\sum_{j=1}^n a_{ij} x_j \right) f_i. \end{aligned}$$

Si representamos $\varphi(x)$ por medio de la matriz columna $\mathcal{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}$ de sus componentes en (f_j) , estas relaciones se pueden condensar en una única relación, que es:

$$(3) \quad \boxed{\mathcal{Y} = M\mathcal{X}}.$$

TEOREMA IX.2.2

|| Sean E, F, G tres K -espacios vectoriales de dimensión finita, de bases respectivas $(e_i)_{1 \leq i \leq n}$, $(f_j)_{1 \leq j \leq p}$, $(g_k)_{1 \leq k \leq q}$ y sean

$$\varphi : E \rightarrow F, \quad \psi : F \rightarrow G$$

|| aplicaciones lineales. Designemos por M_φ , M_ψ y $M_{\psi \circ \varphi}$ las matrices asociadas a estas aplicaciones en las bases anteriores. Entonces se verifica:

$$M_{\psi \circ \varphi} = M_\psi \cdot M_\varphi.$$

Demostración

— Puesto que M_φ es una (p, n) -matriz y M_ψ es una (q, p) -matriz, el producto $M_\psi \cdot M_\varphi$ está bien definido.

— Para todo $x \in E$, designamos por \mathcal{X} la matriz columna asociada a x , por \mathcal{Y} la asociada a $\varphi(x)$, por \mathcal{Z} la asociada a $\psi \circ \varphi(x)$, en las bases dadas. Según (3) se puede escribir:

$$\mathcal{Y} = M_\varphi \cdot \mathcal{X} \quad \mathcal{Z} = M_\psi \mathcal{Y} = M_\psi \cdot M_\varphi \mathcal{X}$$

y $\mathcal{Z} = M_{\psi \circ \varphi} \mathcal{X}$, de donde se deduce: $M_{\psi \circ \varphi} = M_\psi \cdot M_\varphi$. c.q.d.

Con la ayuda de IX.2.1 resulta:

COROLARIO

|| Sea E un K -espacio vectorial de dimensión finita $n \geq 1$, y sea (e_1, \dots, e_n) una base de E . La aplicación que, a todo endomorfismo φ de E , le hace corresponder su matriz M_φ en (e_i) es un isomorfismo de la K -álgebra $\mathcal{L}_K(E)$ en la K -álgebra $\mathcal{M}_n(K)$ (que hace corresponder la matriz unidad I_n a la aplicación idéntica de E).

En particular, los grupos $GL(E)$ y $GL_n(K)$ son isomorfos, y las matrices *invertibles* corresponden a las aplicaciones lineales biyectivas, es decir, a los *automorfismos* de E .

Obsérvese que los isomorfismos puestos de relieve entre $\mathcal{L}_K(E, F)$ y $\mathcal{M}_{p,n}(K)$, o entre $\mathcal{L}_K(E)$ y $\mathcal{M}_n(K)$, dependen esencialmente de las bases elegidas. Ninguno de estos isomorfismos es privilegiado, y veremos más adelante cómo varían cuando se cambian las bases.

Aplicación

El teorema IX.2.2 y su corolario proporcionan un método que, a menudo, resulta cómodo para calcular el producto de matrices, o la inversa de una matriz, considerando las aplicaciones lineales asociadas.

Ejemplos

1) Sea $A = \begin{bmatrix} 1 & \dots & 1 \\ 1 & \dots & 1 \end{bmatrix}$ la matriz cuadrada de orden n cuyos coeficientes son todos iguales a 1.

La transformación lineal asociada $(x_i) \mapsto (y_i)$ está definida por

$$y_i = \sum_{j=1}^n x_j \quad (i = 1, 2, \dots, n);$$

la transformación lineal $(x_i) \mapsto (z_i)$ asociada a A^2 está definida por

$$z_i = \sum_{k=1}^n y_k = n \sum_{j=1}^n x_j.$$

Se tiene, pues, $A^2 = \begin{bmatrix} n & \dots & n \\ n & \dots & n \end{bmatrix} = nA$. Por recurrencia se ve que, para todo entero $p \geq 1$, se verifica $A^p = n^{p-1} A$.

2) Consideremos la matriz cuadrada de orden n :

$$A = \begin{bmatrix} 1 & a & \dots & a \\ a & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ a & \dots & a & 1 \end{bmatrix}.$$

La transformación asociada $(x_i) \mapsto (y_i)$ está definida por las fórmulas:

$$y_i = (1 - a) x_i + a(x_1 + x_2 + \dots + x_n) \quad (1 \leq i \leq n).$$

Hagamos: $s = \sum x_i$ y $S = \sum y_i$. Por adición, se tiene:

$$S = (1 - a) s + nas = [1 + (n - 1) a] s.$$

Si $1 + (n - 1) a \neq 0$, se obtiene: $s = \frac{1}{1 + (n - 1) a} S$, de donde

$$(1 - a) x_i = y_i - as = y_i - \frac{a}{1 + (n - 1) a} (y_1 + y_2 + \dots + y_n).$$

Si además, $1 - a \neq 0$, estas relaciones proporcionan las x_i en función de las y_i :

$$x_i = \frac{y_i}{1 - a} - \frac{a}{(1 - a) [1 + (n - 1) a]} (y_1 + y_2 + \dots + y_n).$$

Luego, cuando $(1 - a) (1 + (n - 1) a) \neq 0$, la matriz A es invertible, y se tiene:

$$A^{-1} = \frac{1}{(1 - a) [1 + (n - 1) a]} \begin{bmatrix} 1 + (n - 2) a & -a & -a & \dots & -a \\ -a & 1 + (n - 2) a & -a & \dots & -a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a & \dots & \dots & \dots & 1 + (n - 2) a \end{bmatrix}.$$

TEOREMA IX.2.3

Sean E, F dos espacios vectoriales de dimensiones n, p y sea (e_1, \dots, e_n) una base de E , y (f_1, \dots, f_p) una base de F .
 Si $\varphi : E \rightarrow F$ es una aplicación lineal, y si A es la matriz de φ en las bases (e_i) y (f_j) , la matriz de la traspuesta ${}^t\varphi$ en las bases duales (e_i^*) y (f_j^*) es tA .

Demostración. e_i^* es la i -ésima forma coordenada $x \mapsto x_i$ en E , y f_j^* es la j -ésima forma coordenada $y \mapsto y_j$ en F (cf. § VIII.5). Escribimos

$$A = [a_{ij}] \quad (1 \leq i \leq n, \quad 1 \leq j \leq p).$$

En las bases $f_j^* = y_j$ ($j = 1, 2, \dots, p$) y $e_i^* = x_i$ ($i = 1, 2, \dots, n$) la traspuesta ${}^t\varphi$ de φ , está definida por las fórmulas de transformación (establecidas antes):

$$(T) \quad y_j = \sum_{i=1}^n a_{ji} x_i \quad (1 \leq j \leq p).$$

Dichas fórmulas se pueden escribir en la forma equivalente:

$${}^t\varphi(f_j^*) = \sum_{i=1}^n a_{ji} e_i^*.$$

La matriz de ${}^t\varphi$ en las bases consideradas es, pues, tA . c.q.d.

Ejemplos de matrices asociadas a aplicaciones lineales

Los espacios \mathbf{R}^n que intervienen a continuación se hallan dotados de su estructura euclídea orientada canónica.

1) En el espacio euclídeo \mathbf{R}^2 o \mathbf{R}^3 , las homotecias de centro en el origen son transformaciones lineales. Las simetrías respecto a los planos que pasan por el origen son lineales.

Tomemos las bases de llegada y de salida idénticas y ortonormales: (\vec{i}, \vec{j}) o $(\vec{i}, \vec{j}, \vec{k})$.

— En \mathbf{R}^2 , la rotación \mathcal{R}_θ de centro O , y ángulo θ está representada por la matriz:

$$R_\theta = \begin{bmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{bmatrix}.$$

Puesto que $\mathcal{R}_\theta \circ \dots \circ \mathcal{R}_\theta$ (n veces) es $\mathcal{R}_{n\theta}$, se tiene:

$$(R_\theta)^n = R_{n\theta} \quad (\text{corolario del teorema IX.5.2}).$$

Como $\mathcal{R}_\theta \circ \mathcal{R}_\psi = \mathcal{R}_{\psi+\theta}$, se tiene con más generalidad $R_\theta \cdot R_\psi = R_{\theta+\psi}$. En particular, R_θ es invertible, y

$$R_\theta^{-1} = R_{-\theta}.$$

— En \mathbf{R}^3 la rotación de eje Oz y ángulo θ se halla representada por la matriz

$$\begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Se pueden hacer consideraciones análogas a las precedentes.

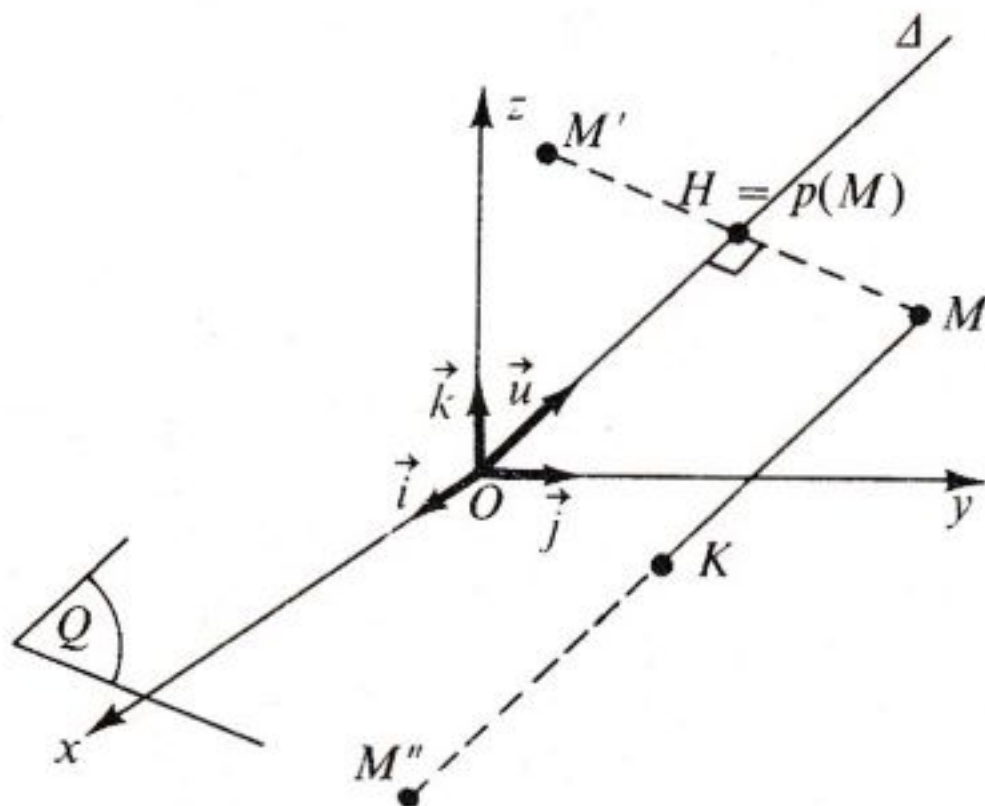
— En \mathbf{R}^3 la simetría respecto del eje Oz está representada por la matriz

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ y la simetría respecto del plano } xOy \text{ está representada por la ma-}$$

$$\text{triz } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

2) Sea Δ la recta de \mathbf{R}^3 , que pasa por el origen, de vector director unitario $\vec{u} \begin{cases} \alpha \\ \beta \\ \gamma \end{cases}$.

Buscamos la matriz de la *proyección ortogonal* p sobre Δ .



Es suficiente calcular $p(\vec{i})$, $p(\vec{j})$, $p(\vec{k})$. Por definición de α , β , γ , se obtiene:

$$p(\vec{i}) = \alpha \vec{u}, \quad p(\vec{j}) = \beta \vec{u}, \quad p(\vec{k}) = \gamma \vec{u},$$

y de ahí (teniendo en cuenta que $\vec{u} = \alpha \vec{i} + \beta \vec{j} + \gamma \vec{k}$), la matriz P de p es:

$$P = \begin{bmatrix} \alpha^2 & \beta\alpha & \gamma\alpha \\ \alpha\beta & \beta^2 & \gamma\beta \\ \alpha\gamma & \beta\gamma & \gamma^2 \end{bmatrix}.$$

Sea M' la simétrica de M respecto de Δ . Puesto que

$$\overrightarrow{OM} + \overrightarrow{OM'} = 2 \overrightarrow{Op(M)},$$

vemos que la matriz de la simetría respecto de Δ es

$$S_{\Delta} = 2P - I_3.$$

Finalmente, si M'' es la simetría de M respecto del plano Q , que pasa por O y es ortogonal a \vec{u} , se tiene:

$$\overrightarrow{OM''} = -\overrightarrow{OM'};$$

de donde la matriz \mathfrak{S}_Q de la simetría respecto del plano Q es: $\mathfrak{S}_Q = I_3 - 2P$.

Designamos por K la proyección ortogonal de M sobre (Q) . Puesto que

$$\overrightarrow{OK} = \frac{1}{2}(\overrightarrow{OM} + \overrightarrow{OM''}),$$

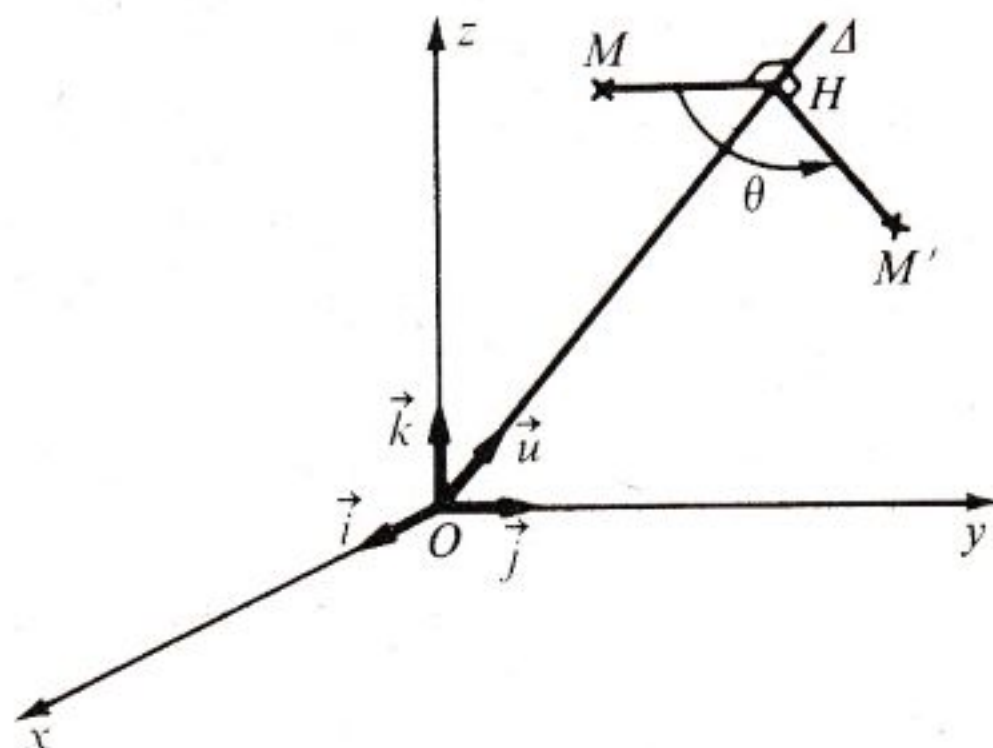
la matriz de la proyección $M \mapsto K$ es $\frac{1}{2}(I_3 + \mathfrak{S}) = I_3 - P$.

3) Sea Δ la recta de \mathbf{R}^3 que pasa por O , de vector unitario $\vec{u} \begin{cases} \alpha \\ \beta \\ \gamma \end{cases}$, y busque-

mos la matriz R de la rotación \mathcal{R} de ángulo θ y eje Δ , en la base canónica $(\vec{i}, \vec{j}, \vec{k})$.

Si H designa la proyección ortogonal de M sobre Δ , tenemos:

$$\overrightarrow{OH} = (\overrightarrow{OM} \cdot \vec{u}) \vec{u}, \quad y \quad \overrightarrow{HM} = (\vec{u} \wedge \overrightarrow{OM}) \wedge \vec{u}.$$



Pongamos $M' = \mathcal{R}(M)$. En la base $(\overrightarrow{HM}, \vec{u} \wedge \overrightarrow{OM}, \vec{u})$ de \mathbf{R}^3 , es posible descomponer $\overrightarrow{OM'}$:

$$\overrightarrow{OM'} = \overrightarrow{OH} + \cos \theta \cdot \overrightarrow{HM} + \operatorname{sen} \theta (\vec{u} \wedge \overrightarrow{OM}),$$

de donde

$$\overrightarrow{OM'} = (\overrightarrow{OM} \cdot \vec{u}) \vec{u} + \cos \theta ((\vec{u} \wedge \overrightarrow{OM}) \wedge \vec{u}) + \operatorname{sen} \theta (\vec{u} \wedge \overrightarrow{OM}).$$

Si en esta fórmula hacemos sucesivamente $\overrightarrow{OM} = \vec{i}, \vec{j}, \vec{k}$, obtenemos la matriz buscada:

$$R = \begin{bmatrix} \alpha^2 + (\beta^2 + \gamma^2) \cos \theta & \alpha\beta - \alpha\beta \cos \theta - \gamma \operatorname{sen} \theta & \alpha\gamma - \alpha\gamma \cos \theta + \beta \operatorname{sen} \theta \\ \alpha\beta - \alpha\beta \cos \theta + \gamma \operatorname{sen} \theta & \beta^2 + (\gamma^2 + \alpha^2) \cos \theta & \beta\gamma - \beta\gamma \cos \theta - \alpha \operatorname{sen} \theta \\ \alpha\gamma - \alpha\gamma \cos \theta - \beta \operatorname{sen} \theta & \beta\gamma - \beta\gamma \cos \theta + \alpha \operatorname{sen} \theta & \gamma^2 + (\alpha^2 + \beta^2) \cos \theta \end{bmatrix}.$$

4) Sea K un cuerpo conmutativo de característica nula y sea \mathcal{P}_n el espacio vectorial de los polinomios $P \in K[X]$, de grado $\leq n$. \mathcal{P}_n es de dimensión $n+1$, y una base de \mathcal{P}_n está formada por los polinomios $1, X, \dots, X^n$. La aplicación $P \mapsto P'$ es una aplicación lineal de \mathcal{P}_n en \mathcal{P}_n y su matriz en la base anterior es:

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & & & n \\ 0 & 0 & \dots & \dots & 0 \end{bmatrix}.$$

La aplicación $P \mapsto Q$ definida por $Q(X) = P(X + a)$ (en donde $a \in K$ es fijo) también es lineal, y su matriz en la base $(1, X, \dots, X^n)$ es la matriz triangular superior

$$M_a = \begin{bmatrix} 1 & a & a^2 & \dots & a^n \\ 0 & 1 & & & \\ \vdots & & \ddots & & \\ \vdots & & & \binom{k}{j} a^{k-j} & \\ \vdots & & & & \ddots \\ 0 & & & & & 1 \end{bmatrix}$$

o sea $M_a = [m_{jk}]$, con $m_{jk} = 0$ si $k < j$ y $m_{jk} = \binom{k}{j} a^{k-j}$ si $k \geq j$. Por la definición, se tiene: $M_a M_b = M_{a+b}$ cualesquiera que sean $a, b \in K$. Esto prueba que las matrices $M_a (a \in K)$ constituyen un grupo multiplicativo isomorfo al grupo aditivo formado por K ; en particular se tiene $(M_a)^{-1} = M_{-a}$ cualquiera que sea $a \in K$ (cf. Ejercicio IX.8).

* Matrices en un anillo conmutativo

Todo lo que acabamos de exponer a partir del § 1 se aplica sin restricciones si se reemplaza el cuerpo K por un anillo conmutativo unífero A , y la expresión «espacio vectorial sobre K » por la de « A -módulo». En el ejemplo 4) del cálculo con matrices, es suficiente suponer que la relación $2u = 0$ implica $u = 0$ en A . (cf. p. 325).

En el § 2, definición IX.2.1 es preciso reemplazar E y F por A -módulos con bases finitas. Según hemos señalado anteriormente, estos módulos poseen una dimensión bien definida (todas las bases tienen el mismo número de elementos). Las definiciones y propiedades que siguen a la definición IX.2.1 subsisten entonces sin variaciones.

§ IX.3 CAMBIO DE BASE

● Sea E un espacio vectorial de dimensión n , y sea (e_1, e_2, \dots, e_n) una base de E . Si v_1, \dots, v_n son elementos de E , la **matriz P de los v_i en la base (e_i)** es, por definición, la matriz de la aplicación lineal φ tal que

$$\varphi(e_i) = v_i.$$

φ es biyectiva si, y sólo si, v_1, \dots, v_n son libres, o lo equivalente, cuando la matriz P es **invertible**. Cuando ocurre así, se dice que P es la **matriz de cambio** de las (e_i) a las (v_i) .

Cambio de coordenadas

Sean $(e_1, \dots, e_n), (f_1, \dots, f_n)$ dos bases del espacio vectorial E , y sea P la matriz de cambio de las (e_i) a las (f_j) , con $P = [p_{ij}]_{1 \leq i, j \leq n}$.

Sea $x \in E$, y designemos por \mathcal{X} (resp. \mathcal{Y}) a la matriz columna de las coordenadas de x en la base (e_i) (resp. en la base (f_j)):

$$\mathcal{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \mathcal{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Por definición, las relaciones siguientes son verdaderas:

$$f_j = \sum_{i=1}^n p_{ij} e_i \quad (1 \leq j \leq n), \quad x = \sum_{i=1}^n x_i e_i = \sum_{j=1}^n y_j f_j.$$

Se deduce:

$$\sum_{j=1}^n y_j f_j = \sum_{j=1}^n \sum_{i=1}^n p_{ij} y_j e_i = \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij} y_j \right) e_i.$$

De donde:

$$x_i = \sum_{j=1}^n p_{ij} y_j \quad 1 \leq i \leq n.$$

Estas fórmulas se pueden condensar en la fórmula matricial siguiente, llamada *fórmula del cambio de coordenadas*:

(1)

$$\boxed{\mathcal{X} = P\mathcal{Y}}.$$

Problema del cambio de base

Sean E, F dos espacios vectoriales, y sea $\varphi : E \rightarrow F$ una aplicación lineal. Designamos por

- $(e_1, \dots, e_n), (e'_1, \dots, e'_n)$ dos bases de E ;
- $(f_1, \dots, f_p), (f'_1, \dots, f'_p)$ dos bases de F ;
- P la matriz de cambio de las (e_i) a las (e'_j) ;
- Q la matriz de cambio de las (f_i) a las (f'_j) ;
- A la matriz de φ en las bases (e_i) y (f_j) ;
- B la matriz de φ en las bases (e'_i) y (f'_j) ;

- x un elemento de E , $y = \varphi(x)$ su imagen por φ ;
- \mathcal{X} y \mathcal{X}' las matrices columnas de las coordenadas de x en (e_i) y (e'_i) ;
- \mathcal{Y} e \mathcal{Y}' las matrices columnas de las coordenadas de y en (f_j) y (f'_j) .

Vamos a establecer una relación que ligue P , Q , A , B . Para ello, aplicamos (1) y la fórmula (3) del § 2:

$$(2) \quad \mathcal{X} = P\mathcal{X}' \quad \mathcal{Y} = Q\mathcal{Y}'$$

$$(3) \quad \mathcal{Y} = A\mathcal{X} \quad \mathcal{Y}' = B\mathcal{X}'$$

(2) nos da, llevándola a la primera de las fórmulas de (3):

$$Q\mathcal{Y}' = AP\mathcal{X}',$$

o, puesto que Q es invertible:

$$\mathcal{Y}' = Q^{-1} AP\mathcal{X}',$$

de donde resulta la relación fundamental:

$$(4) \quad \boxed{B = Q^{-1} AP}.$$

Caso particular: $E = F$, y $e_i = f_i$, $e'_i = f'_i$ para $1 \leq i \leq n$, luego $P = Q$. La fórmula (4) se escribe en este caso:

$$(5) \quad \boxed{B = P^{-1} AP},$$

en donde todas las matrices son cuadradas de orden n .

DEFINICIÓN IX.3.1

*Dos (p, n) -matrices A y B se llaman **equivalentes** si existe una matriz cuadrada invertible Q de orden p , y una matriz cuadrada invertible P de orden n , tales que*

$$B = QAP.$$

Se comprueba sin ninguna dificultad que la relación « A y B son equivalentes» es una relación de equivalencia en $M_{p,n}(K)$. Ante todo definiremos el *rango de una (p, n) -matriz A* . Para ello probaremos que, si E y F son dos K -espacios vectoriales

de dimensiones respectivas n y p , provistos de las bases $\beta = (e_1, \dots, e_n)$ y $\gamma = (f_1, \dots, f_p)$, y si $\varphi = \varphi_{\beta, \gamma}$ es la aplicación lineal de E en F cuya matriz en las bases β y γ es A , entonces el rango de $\varphi_{\beta, \gamma}$ depende sólo de A . En efecto, sean E' y F' otros dos espacios vectoriales, de bases respectivas $\beta' = (e'_1, \dots, e'_n)$ y $\gamma' = (f'_1, \dots, f'_p)$ y sea $\varphi' = \varphi_{\beta', \gamma'}$. Designamos por u (resp. v) al isomorfismo de E en E' (resp. de F en F') tal que $u(e_i) = e'_i$ para todo i (resp. $v(f_j) = f'_j$ para todo j).

Entonces se comprueba fácilmente que

$$\varphi' = v \circ \varphi \circ u^{-1}$$

luego las dimensiones de $\text{Im}(\varphi)$ e $\text{Im}(\varphi')$ son iguales, y de ahí nuestra afirmación. Establecido esto, por definición, el *rango de A* es el rango común a todas las $\varphi_{\beta, \gamma}$. Se le designa por $\text{rg}(A)$.

Sea entonces $\varphi : K^n \rightarrow K^p$ una aplicación lineal que tenga a A por matriz en una cierta elección de bases de K^n y K^p , y sean $r = \text{rg}(A)$, $N = \text{Ker}(\varphi)$, $I = \text{Im}(\varphi)$. Se puede hallar una base de K^n , por ejemplo (e_1, \dots, e_n) , tal que (e_{r+1}, \dots, e_n) sea una base de N ; entonces $(f_1 = \varphi(e_1), \dots, f_r = \varphi(e_r))$ es una base de I , y (ver p. 293) existen $f_{r+1} \in F, \dots, f_p \in F$ tales que (f_1, \dots, f_p) es una base de F . En las bases (e_1, \dots, e_n) y (f_1, \dots, f_p) así definidas, la matriz de φ se escribe:

$$(6) \quad \left\{ \begin{array}{cccc} 1 & \dots & 0 & \\ 0 & 1 & 0 & 0 \\ \vdots & \ddots & \vdots & \\ 0 & \dots & 1 & \dots \\ & 0 & \vdots & 0 \end{array} \right\} \begin{array}{c} r \\ p \end{array} \quad \underbrace{\hspace{10em}}_n$$

Teniendo en cuenta (4), vemos que ha quedado demostrado el siguiente:

TEOREMA IX.3.1

|| Toda (p, n) -matriz de rango r es equivalente a la matriz (6).

DEFINICIÓN IX.3.2

⎵ Dos (n, n) -matrices A y B son **semejantes** si existe una matriz cuadrada invertible P de orden n tal que

$$B = P^{-1} A P.$$

La relación « A y B son semejantes» es una relación de equivalencia en $M_n(K)$, más fina que « A y B son equivalentes».

El lenguaje de los *grupos que operan sobre un conjunto* (cf. Cap. II, § 8) permite precisar esta relación de equivalencia. Para toda matriz $P \in GL(n, K)$ y toda matriz $A \in M_n(K)$, establecemos:

$$(7) \quad A * P = P^{-1} A P.$$

Se comprueba inmediatamente que:

$$A * I_n = A \quad \text{y} \quad (A * P) * Q = A * (PQ).$$

Luego, el grupo $GL(n, K)$ opera por la derecha sobre $M_n(K)$ por medio de la fórmula (7).

Las clases de equivalencia de $M_n(K)$ por la relación « A y B son semejantes» son precisamente las *órbitas* de $M_n(K)$ según $GL(n, K)$.

Si $A \in GL(n, K)$, $A * P \in GL(n, K)$. Con otras palabras, $GL(n, K)$ es *reunión de órbitas*. Las órbitas de $GL(n, K)$ son precisamente las *clases de conjugación* de los elementos de $GL(n, K)$ (cf. Cap. II, § 8, ejemplo 3).

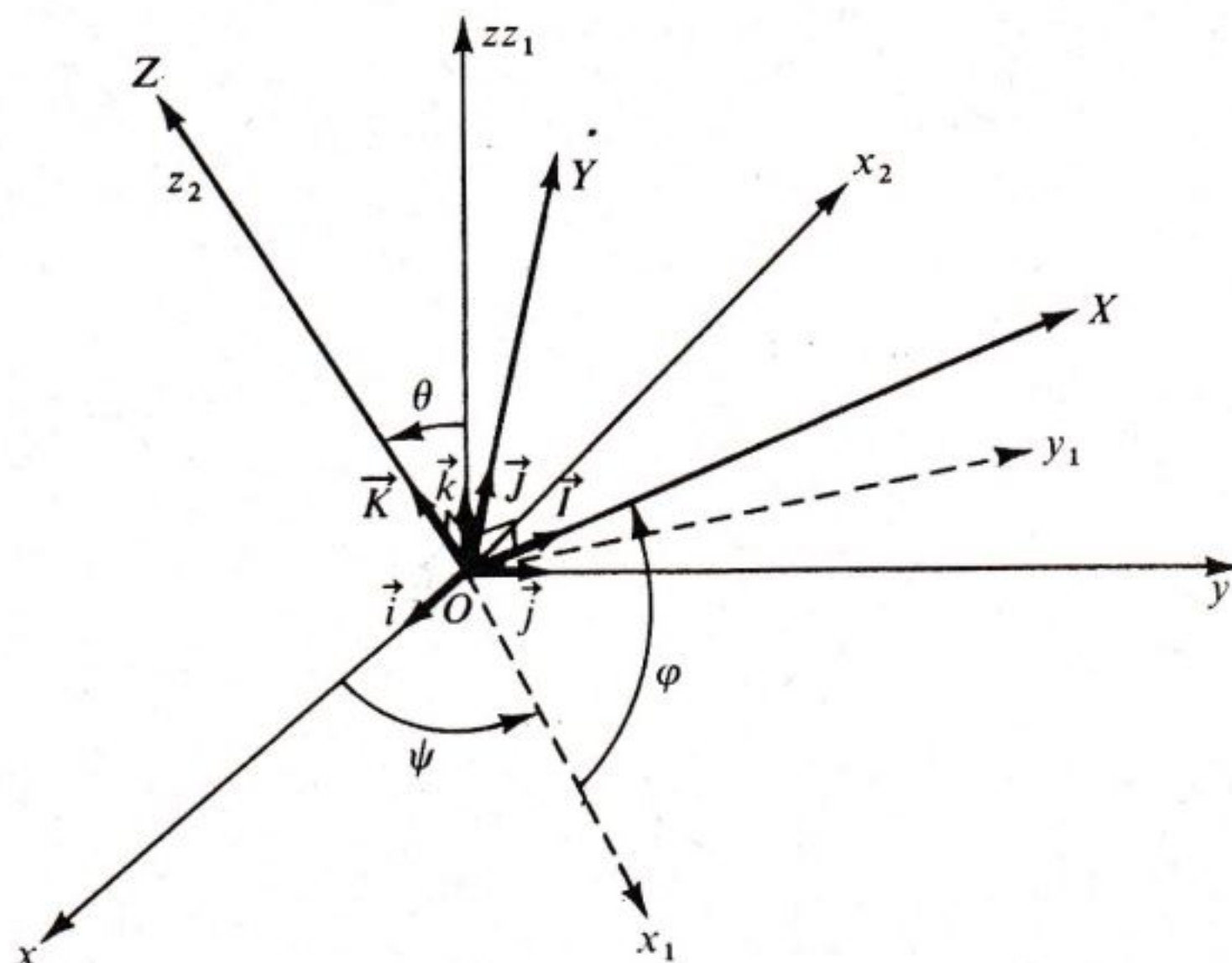
El teorema IX.6.1 permite *clasificar* las (p, n) -matrices salvo equivalencias. Se comprueba que «lo que clasifica» es el *rango* (dos matrices son equivalentes si, y sólo si, tienen el mismo rango).

El problema de la clasificación de las matrices cuadradas salvo semejanzas equivale a la caracterización de las órbitas de $M_n(K)$. En el capítulo XI lo resolveremos en el caso en que el cuerpo K sea algebraicamente cerrado. (Reducida de Jordan.) Con mayor precisión, demostraremos la existencia, en cada órbita, de matrices particulares, llamadas de Jordan, cuyo tipo caracteriza la órbita en cuestión, y éste es el problema completo de la *reducción* de las matrices cuadradas.

Ejemplo de matriz de cambio: ángulos de Euler

Sea $Oxyz$ una referencia ortonormal de \mathbf{R}^3 , $OXYZ$ otra referencia ortonormal con el mismo origen. Se determina $OXYZ$ con la ayuda de los tres *ángulos de Euler*:

- El plano OXY corta a Oxy según una recta, que orientaremos según Ox_1 .
- Al ángulo $\psi = (\overrightarrow{Ox}, \overrightarrow{Ox_1})$ se le llama *precesión*. La orientación de Ox_1 define una orientación en todo el plano ortogonal a Ox_1 , en particular sobre OzZ .
- Al ángulo $\theta = (\overrightarrow{Oz}, \overrightarrow{OZ})$ se le llama *nutación*.
- Finalmente, al ángulo $\varphi = (\overrightarrow{Ox_1}, \overrightarrow{OX})$, medido en el plano orientado OXY , se le llama *rotación propia*.



Vamos a calcular la matriz de cambio A de $(\vec{i}, \vec{j}, \vec{k})$ en $(\vec{I}, \vec{J}, \vec{K})$ (vectores unitarios respectivamente de $Oxyz$ y $OXYZ$).

Para ello introducimos las referencias intermedias $Ox_1y_1z_1$ y $Ox_2y_2z_2$ ortonormadas y directas tales que

$$Oz_1 = Oz, \quad Ox_2 = Ox_1, \quad Oz_2 = OZ.$$

Matriz de cambio de $(Oxyz)$ a $(Ox_1y_1z_1)$: $M = \begin{bmatrix} \cos \psi & -\text{sen } \psi & 0 \\ \text{sen } \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Matriz de cambio de $(Ox_1y_1z_1)$ a $(Ox_2y_2z_2)$: $N = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\text{sen } \theta \\ 0 & \text{sen } \theta & \cos \theta \end{bmatrix}$

Matriz de cambio de $(Ox_2y_2z_2)$ a $(OXYZ)$: $P = \begin{bmatrix} \cos \varphi & -\text{sen } \varphi & 0 \\ \text{sen } \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$

De donde: $A = MNP$.

Sea ahora \mathcal{R} la rotación de eje OZ y de ángulo α , y busquemos su matriz R en $(\vec{i}, \vec{j}, \vec{k})$.

La matriz de \mathcal{R} en $(Ox_2y_2z_2)$ es evidentemente

$$Q = \begin{bmatrix} \cos \alpha & -\operatorname{sen} \alpha & 0 \\ \operatorname{sen} \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Puesto que la matriz de cambio de $(Oxyz)$ a $(Ox_2y_2z_2)$ es MN , se deduce, con la ayuda de (5), que:

$$Q = N^{-1} M^{-1} R M N, \text{ de donde } R = M N Q N^{-1} M^{-1}$$

que constituye una nueva forma de la matriz de rotación más general (cf. ejemplo § IX.1).

Nota. Para cada referencia ortonormal $OXYZ$, existen dos posibles elecciones de la orientación de Ox_1 . Además, la recta Ox_1 está indeterminada cuando $Oz = OZ$. Existe, pues, cierta dificultad para definir los ángulos de Euler como *funciones diferenciables* de la referencia $OXYZ$. En la práctica, se intenta definirlos «por continuidad», a partir de una elección relativa a una posición particular de $OXYZ$, lo que equivale a imponer a ψ , θ , φ que sean funciones continuas de los «parámetros» de los que depende la referencia. Pero la definición rigurosa de estas funciones sobrepasa los límites de esta obra.

Capítulo X

Los determinantes y sus aplicaciones

- En este capítulo, el cuerpo de base K se supone conmutativo.

§ X.1 APLICACIONES Y FORMAS MULTILINEALES

DEFINICIÓN X.1.1

Sean E_1, \dots, E_p y F espacios vectoriales sobre K (no necesariamente de dimensión finita). Una aplicación

$$\varphi : E_1 \times E_2 \times \dots \times E_p \rightarrow F$$

se denomina ***p*-lineal** si, para todo índice i , y para todo sistema de elementos $x_j \in E_j$ ($j \neq i$), la aplicación parcial:

$$x_i \mapsto \varphi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n),$$

de E_i en F , es lineal.

Si $p = 1$ se obtiene la noción de aplicación lineal. Si $p = 2$, se obtiene la noción de aplicación *bilineal*.

Cuando $F = K$, se tiene la noción de *forma p-lineal* en $E_1 \times E_2 \times \dots \times E_p$. A una forma *p*-lineal en E^p se le llama simplemente una *forma p-lineal en E*.

Las propiedades que siguen son inmediatas:

— Si φ es una aplicación *p*-lineal de $E_1 \times E_2 \times \dots \times E_p$ en F , φ toma el valor 0 para toda *n*-pla (x_1, x_2, \dots, x_n) en la que *uno* de los x_i sea nulo.

— El conjunto de las aplicaciones p -lineales de $E_1 \times E_2 \times \dots \times E_p$ en F es un subespacio vectorial del espacio vectorial de las aplicaciones de $E_1 \times E_2 \times \dots \times E_p$ en F . A este subespacio se le designa por $\mathcal{L}_p(E_1, \dots, E_p; F)$; al espacio de las formas p -lineales en E se le designa simplemente por $\mathcal{L}_p(E)$.

— Si F es suma directa de F_1, \dots, F_k , el espacio $\mathcal{L}_p(E_1, E_2, \dots, E_p; F)$ es isomorfo a la suma directa de los espacios $\mathcal{L}_p(E_1, \dots, E_p, F_m)$ ($1 \leq m \leq k$). Se obtiene un isomorfismo asociando a cada

$$\varphi \in \mathcal{L}_p(E_1, \dots, E_p; F),$$

la k -pla $(\varphi_m)_{1 \leq m \leq k}$, en donde φ_m es la aplicación p -lineal en $E_1 \times E_2 \times \dots \times E_p$ definida por $\varphi_m = p_m \circ \varphi$ (en donde p_m designa la proyección de F en F_m).

A las aplicaciones φ_m se les llama *componentes* de φ .

En particular, si F es de dimensión finita, vemos que el estudio de las aplicaciones p -lineales: $E_1 \times \dots \times E_p \rightarrow F$ se reduce al de las formas p -lineales en $E_1 \times \dots \times E_p$.

— Si E_1, E_2, \dots, E_p y F son de dimensión finita, también lo es $\mathcal{L}_p(E_1, \dots, E_p; F)$. En efecto, según la propiedad anterior, es suficiente demostrarlo cuando $F = K$. A tal fin, consideremos, para todo i , una base $(e_{i,1}, e_{i,2}, \dots, e_{i,r_i})$ de E_i , y sea $\varphi \in \mathcal{L}_p(E_1, \dots, E_p; K)$. Para todo $x_i \in E_i$, ponemos:

$$x_i = \sum_{j=1}^{r_i} a_{ji} e_{i,j}.$$

Desarrollando $\varphi(x_1, \dots, x_p)$, por la p -linealidad, se obtiene:

$$\varphi(x_1, x_2, \dots, x_p) = \sum_{\substack{1 \leq j_1 \leq r_1 \\ 1 \leq j_2 \leq r_2 \\ \vdots \\ 1 \leq j_p \leq r_p}} a_{j_1,1} a_{j_2,2} \dots a_{j_p,p} \varphi(e_{1,j_1}, e_{2,j_2}, \dots, e_{p,j_p}).$$

[Para que se comprenda perfectamente el segundo miembro, tomemos, por ejemplo, $p = 2$, $r_1 = r_2 = 2$:

$$x_1 = a_{1,1} e_{1,1} + a_{2,1} e_{1,2} \quad x_2 = a_{1,2} e_{2,1} + a_{2,2} e_{2,2},$$

$$\varphi(x_1, x_2) = \varphi(a_{1,1} e_{1,1} + a_{2,1} e_{1,2}, a_{1,2} e_{2,1} + a_{2,2} e_{2,2})$$

$$= a_{11} a_{12} \varphi(e_{11}, e_{2,1}) + a_{11} a_{22} \varphi(e_{11}, e_{22})$$

$$+ a_{21} a_{12} \varphi(e_{12}, e_{2,1}) + a_{21} a_{22} \varphi(e_{12}, e_{22}).]$$

Recíprocamente, toda aplicación ψ de $E_1 \times E_2 \times \dots \times E_p$ en K , definida por una fórmula del tipo

$$(P) \quad \psi(x_1, x_2, \dots, x_p) = \sum_{\substack{1 \leq j_1 \leq r_1 \\ \vdots \\ 1 \leq j_p \leq r_p}} A_{j_1 \dots j_p} a_{j_1, 1} a_{j_2, 2} \dots a_{j_p, p},$$

en donde $A_{j_1 j_2 \dots j_p} \in K$ para todo $(j_1, \dots, j_p) \in \mathbf{N}_{r_1}^* \times \mathbf{N}_{r_2}^* \times \dots \times \mathbf{N}_{r_p}^*$, es una forma p -lineal en $E_1 \times \dots \times E_p$. Se deduce sin ninguna dificultad que el espacio

$$\mathcal{L}_p(E_1, \dots, E_p; K)$$

es de dimensión r_1, r_2, \dots, r_p (en donde $r_i = \dim(E_i)$). c.q.d.

Los subespacios más importantes de $\mathcal{L}_p(E)$ son:

— el espacio de las formas p -lineales *simétricas*, que estudiaremos más adelante en un caso particular (cf. Cap. XII);

— el espacio de las formas p -lineales *alternadas*, que estudiaremos en este capítulo. Estas formas sirven para estudiar el rango de un sistema de vectores de E .

DEFINICIÓN X.1.2

Sea E un espacio vectorial. Una forma φ , p -lineal sobre E , es **simétrica** si, para toda p -pla $(x_i)_{1 \leq i \leq p}$ ($x_i \in E$) y toda permutación $\sigma \in \mathfrak{S}_p$, se tiene:

$$\varphi(x_1, \dots, x_p) = \varphi(x_{\sigma(1)} \dots x_{\sigma(p)}).$$

DEFINICIÓN X.1.3

Una forma p -lineal φ en un espacio vectorial E se llama **alternada** si se tiene $\varphi(x_1, x_2, \dots, x_p) = 0$ cada vez que las $(x_i)_{1 \leq i \leq p}$ no son todas distintas.

Es inmediato que el conjunto de las formas p -lineales alternadas en E forman un subespacio de $\mathcal{L}_p(E)$. Este subespacio lo designaremos por:

$$\Lambda^{*p}(E).$$

Para toda forma p -lineal φ en E , y toda permutación $\sigma \in \mathfrak{S}_p$, designaremos por medio de $\sigma^*(\varphi)$ a la forma p -lineal definida por:

$$\sigma^*(\varphi)(x_1, \dots, x_p) = \varphi(x_{\sigma(1)}, \dots, x_{\sigma(p)}).$$

Se tiene: $\forall \rho \in \mathfrak{S}_p, \forall \sigma \in \mathfrak{S}_p, \forall \varphi \in \Lambda^{*p}(E)$

$$\rho^* \cdot (\sigma^*(\varphi)) = (\rho\sigma)^*(\varphi), \quad \text{y} \quad e^*(\varphi) = \varphi \quad (e: \text{permutación idéntica})^{(1)}.$$

TEOREMA X.1.1

|| Sea φ una forma p -lineal en el espacio vectorial E . Para que φ sea alternada, es necesario que, para toda trasposición τ de \mathbf{N}_p^* , se verifique:

(1) $\tau^*(\varphi) = -\varphi$;

|| y esta condición es suficiente cuando K no tiene característica 2.

Demostración

— Sean j y k enteros $\leq p, j < k$, y sea τ la trasposición que cambia j y k . Si (1) se verifica, y si $(x_i)_{1 \leq i \leq p}$ es una familia de elementos de E tal que $x_j = x_k$, se tiene

$$\tau^*(\varphi)(x_1, \dots, x_p) = -\varphi(x_1, \dots, x_p) \quad \text{según (1),}$$

y
$$\tau^*(\varphi)(x_1, \dots, x_p) = \varphi(x_1, \dots, x_p),$$

de donde

$$2\varphi(x_1, \dots, x_p) = 0, \quad \text{o sea} \quad \varphi(x_1, \dots, x_p) = 0$$

cuando K no tiene característica 2.

En este caso la condición (1) es suficiente.

— Si φ es alternada, se puede escribir (ahora el cuerpo K es cualquiera):

$$\begin{aligned} 0 &= \varphi(x_1, \dots, x_j + x_k, \dots, x_j + x_k, \dots, x_p) = \varphi(x_1, \dots, x_p) + \tau^*(\varphi)(x_1, \dots, x_p) + \\ &\quad + \varphi(x_1, \dots, x_j, \dots, x_j, \dots, x_p) + \varphi(x_1, \dots, x_k, \dots, x_k, \dots, x_p) = \\ &= \varphi(x_1, \dots, x_p) + \tau^*(\varphi)(x_1, \dots, x_p). \end{aligned}$$

La condición (1) es, pues, necesaria. ||

COROLARIO

|| Si φ es una forma p -lineal alternada sobre E , para toda permutación $\sigma \in \mathfrak{S}_p$, se tiene:

||
$$\sigma^*(\varphi) = \varepsilon(\sigma) \varphi \quad (\varepsilon(\sigma) \text{ designa la signatura de } \sigma).$$

⁽¹⁾ En otras palabras $(\sigma, \varphi) \mapsto \sigma^*(\varphi)$ define una operación por la izquierda de \mathfrak{S}_p en $\Lambda^{*p}(E)$.

En efecto, para verlo es suficiente descomponer σ en producto de trasposiciones τ_1, \dots, τ_k , y entonces se tiene $\varepsilon(\sigma) = (-1)^k$, y aplicando la fórmula (1), se obtiene el resultado deseado. c.q.d.

En particular, vemos que una forma p -lineal alternada es *invariante respecto del grupo alternado* de orden p .

Nota. Una forma p -lineal φ en E se llama *antisimétrica* si verifica:

$$\forall \sigma \in \mathfrak{S}_p, \quad \sigma^*(\varphi) = \varepsilon(\sigma) \varphi.$$

Si K es de característica $\neq 2$, existe identidad entre las nociones de forma alternada y antisimétrica. Si K es de característica 2, toda forma alternada es antisimétrica, pero el recíproco es falso.

TEOREMA X.1.2

|| Sea E un espacio vectorial que posea una base finita (e_1, \dots, e_n) :
 a) para todo entero $k > n$, se tiene: $\wedge^{*k}(E) = \{0\}$;
 b) el espacio $\wedge^{*n}(E)$ tiene dimensión 1.

Demostración

● a) Sean $(x_i)_{1 \leq i \leq k}$ elementos de E , y pongamos $x_i = \sum_{j=1}^n a_{ij} e_j$. Desarrollemos $\varphi(x_1, \dots, x_k)$ teniendo en cuenta la k -linealidad ($\varphi \in \wedge^{*k}(E)$). Se obtiene:

$$(2) \quad \varphi(x_1, \dots, x_k) = \sum_{\chi \in \mathcal{F}_{k,n}} a_{\chi(1),1} a_{\chi(2),2} \dots a_{\chi(k),k} \varphi(e_{\chi(1)}, \dots, e_{\chi(k)})$$

designando por $\mathcal{F}_{k,n}$ al conjunto de las aplicaciones de \mathbf{N}_k^* en \mathbf{N}_n^* .

Si $k > n$, para cada $\chi \in \mathcal{F}_{k,n}$ dos, por lo menos, de los $e_{\chi(i)}$ son iguales. Puesto que φ es alternada, (2) prueba que en este caso, $\varphi(x_1, \dots, x_k) = 0$.

b) Si $k = n$ la fórmula (2) nos da la expresión que debe tener una aplicación $\varphi \in \wedge^{*n}(E)$:

$$\varphi(x_1, \dots, x_n) = \sum_{\chi \in \mathcal{F}_{n,n}} a_{\chi(1),1} \dots a_{\chi(n),n} \varphi(e_{\chi(1)}, \dots, e_{\chi(n)});$$

en donde los términos que corresponden a las aplicaciones χ no biyectivas son nulos. Esta fórmula se reduce, pues, a:

$$(3) \quad \varphi(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1),1} \dots a_{\sigma(n),n} \varphi(e_{\sigma(1)}, \dots, e_{\sigma(n)})$$

la cual se escribe, si tenemos en cuenta: $\sigma^*(\varphi) = \varepsilon(\sigma) \varphi$:

$$(4) \quad \varphi(x_1, \dots, x_n) = \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} \right) \varphi(e_1, \dots, e_n).$$

Luego φ es un múltiplo de la función Δ definida por

$$\Delta(x_1, x_2, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n},$$

con lo que queda demostrado que $\dim \wedge^{*n}(E) \leq 1$.

Recíprocamente, la función Δ es una forma n -lineal alternada *no nula* en E .

Es evidente que Δ es una forma n -lineal (fórmula (P)), y probemos que es alternada. Para ello designemos por τ a la trasposición que intercambia i y j ($1 \leq i < j \leq n$). Vamos a probar que si $x_i = x_j$, se tiene $\Delta(x_1, x_2, \dots, x_n) = 0$.

Si designamos por \mathcal{A}_n al grupo de las permutaciones pares (grupo alternado), tenemos evidentemente:

$$\Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathcal{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} + \sum_{\sigma \in \mathfrak{S}_n \setminus \mathcal{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}.$$

La aplicación $\sigma \mapsto \sigma \circ \tau$ es una biyección de \mathcal{A}_n en $\mathfrak{S}_n \setminus \mathcal{A}_n$, luego podemos escribir:

$$(5) \quad \Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathcal{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} + \sum_{\sigma \in \mathcal{A}_n} \varepsilon(\sigma\tau) a_{\sigma\tau(1),1} \dots a_{\sigma\tau(n),n},$$

es decir (teniendo en cuenta que $\varepsilon(\sigma\tau) = -\varepsilon(\sigma)$):

$$(6) \quad \Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathcal{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} - \sum_{\sigma \in \mathcal{A}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(j),i} \dots a_{\sigma(i),j} \dots a_{\sigma(n),n}.$$

Luego, puesto que $x_i = x_j$, se tiene $a_{k,i} = a_{k,j}$ para todo $k = 1, 2, \dots, n$, de donde,

$$a_{\sigma(j),i} = a_{\sigma(j),j} \quad \text{y} \quad a_{\sigma(i),j} = a_{\sigma(i),i}$$

y la relación (6) implica $\Delta(x_1, \dots, x_n) = 0$.

Finalmente, comprobamos que $\Delta \neq 0$. En efecto, se tiene: $e_i = \sum_j \delta_{ji} e_j$, de donde

$$\Delta(e_1, \dots, e_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \delta_{\sigma(1),1} \dots \delta_{\sigma(n),n} = \delta_{1,1} \dots \delta_{n,n} = 1. \text{ c.q.d.}$$

La demostración de la afirmación *b)* ha consistido en establecer la siguiente proposición, que le es equivalente, y que, sin embargo, resultará útil enunciarla:

X.1.3 Si (e_1, \dots, e_n) es una base del espacio vectorial E , existe una forma n -lineal **alternada** Δ en E , **única** que verifica $\Delta(e_1, \dots, e_n) = 1$.

DEFINICIÓN X.1.4

Con las notaciones de X.1.3, el **determinante de n vectores** x_1, \dots, x_n de E , respecto de la base (e_1, \dots, e_n) es el escalar

$$\Delta(x_1, \dots, x_n).$$

Si hacemos $x_i = \sum_{j=1}^n a_{ji} e_j$, se tiene:

$$\Delta(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}.$$

Designemos por $\det_{\mathcal{B}}(x_1, \dots, x_n)$ al determinante de n vectores x_1, \dots, x_n en una base \mathcal{B} . Para todas las bases $\mathcal{B} = (e_1, \dots, e_n)$, $\mathcal{C} = (f_1, \dots, f_n)$ de E , y para todo sistema de vectores x_1, \dots, x_n de E , se tiene entonces («fórmula de Chasles»):

$$(3) \quad \det_{\mathcal{B}}(x_1, \dots, x_n) = \det_{\mathcal{C}}(x_1, \dots, x_n) \times \det_{\mathcal{B}}(f_1, f_2, \dots, f_n).$$

En efecto, si hacemos variar arbitrariamente los x_i , ambos miembros son formas n -lineales alternadas en E , que toman el mismo valor para el sistema $x_1 = f_1, \dots, x_n = f_n$, luego son iguales.

Más adelante volveremos sobre esta noción (§ 2).

Extensiones. La demostración de X.1.2 *a)* prueba que, si k elementos $x_i \in E$ son combinaciones lineales de p elementos $e_j \in E$, y si $p < k$, entonces, para toda forma k -lineal alternada φ en E , se tiene: $\varphi(x_1, \dots, x_k) = 0$. En particular, esto se verifica cuando x_1, \dots, x_k están ligados, de donde:

X.1.4 Si k elementos x_1, \dots, x_k de un espacio vectorial E están ligados, entonces
 \parallel toda forma φ , k -lineal y alternada en E , verifica: $\varphi(x_1, \dots, x_k) = 0$.

X.1.5 Sea B una base cualquiera del espacio vectorial E . Si e_1, \dots, e_p son elementos de B , distintos ⁽¹⁾, existe una forma φ , p -lineal y alternada en E , tal que $\varphi(e_1, \dots, e_p) = 1$.

⁽¹⁾ Si E es de dimensión finita, dar B es superfluo; es suficiente suponer que la familia (e_1, \dots, e_p) es libre.

Demostración. Sea

$$F = \text{Vect}(\{e_1, \dots, e_p\}) \quad \text{y} \quad G = \text{Vect}(B \setminus \{e_1, \dots, e_p\});$$

E es suma directa de F y G . Designemos por f la proyección de E en F paralelamente a G . Según el teorema X.1.2, existe una forma p -lineal y alternada φ_1 en F tal que $\varphi_1(e_1, \dots, e_p) = 1$, puesto que (e_1, \dots, e_p) es una base de F .

La aplicación

$$\varphi : (x_1, \dots, x_p) \mapsto \varphi_1(f(x_1), f(x_2), \dots, f(x_p))$$

es entonces una forma p -lineal alternada en E que responde a nuestros propósitos. \square

* Observemos que *todo lo que hemos realizado es rigurosamente independiente del teorema de la dimensión*, y sólo precisa de la definición y de las propiedades elementales de una base. Vamos a deducir una nueva demostración del teorema VIII.3.3 si K es conmutativo.

COROLARIO

\parallel Si E es un espacio vectorial con una base finita B de n elementos, cualquier otra base de E es finita y tiene n elementos.

Demostración. Sea C otra base de E . Si $k = \text{card}(C)$ fuera $> n$ ⁽¹⁾, la proposición X.1.5 implicaría la existencia de una forma $(n+1)$ -lineal, alternada, no nula φ , lo que estaría en contradicción con el teorema X.1.2 aplicado a la base B . Intercambiando B y C veríamos igualmente que tampoco se verifica $k < n$. Por lo tanto se tiene $k = n$. c.q.d.

Nota. Todo lo que hemos realizado en este §, con excepción de X.1.4, se puede repetir, palabra por palabra, en el ámbito de los *módulos sobre un anillo conmutativo* A . Las definiciones son las mismas, y todos los teoremas demostrados en este § son válidos, con la excepción de X.1.4 y de la segunda parte de X.1.1. En particular, se obtiene una generalización considerable del teorema de la dimensión:

TEOREMA

\parallel Sea A un anillo conmutativo unífero y sea M un A -módulo. Si M posee una base finita de n elementos, cualquier otra base de M es finita y posee n elementos.

⁽¹⁾ Esta hipótesis engloba el caso en que C fuese infinito.

Si K es conmutativo, esta demostración del teorema de la dimensión resulta más general que la dada en el capítulo IX. Sin embargo, la del capítulo IX está mucho más adaptada al ámbito de los espacios vectoriales, y permite desarrollar rápidamente consecuencias que únicamente son válidas en un espacio vectorial (por ejemplo, el teorema de la base incompleta).

§ X.2 DETERMINANTES

● A partir de este momento, sólo consideraremos espacios vectoriales de dimensión finita ≥ 1 sobre el cuerpo conmutativo K .

Determinante de un endomorfismo

Sea E un espacio vectorial de dimensión n , y sea φ una forma n -lineal alternada no nula en E . Por otro lado, sea u un endomorfismo de E .

Para toda n -pla de elementos $x_i \in E$ ($1 \leq i \leq n$), ponemos

$$\varphi_u(x_1, \dots, x_n) = \varphi(u(x_1), u(x_2), \dots, u(x_n)).$$

φ_u es evidentemente una forma n -lineal alternada. Luego (T. X.1.2) φ_u es un múltiplo de φ . Vamos a ver que el factor de proporcionalidad sólo depende de φ . En efecto, si ψ es otra forma n -lineal alternada no nula, existe un $\rho \in K$ tal que $\psi = \rho\varphi$ (T. X.1.2). Es evidente que

$$\psi_u = (\rho\varphi)_u = \rho \cdot \varphi_u,$$

luego, si $\varphi_u = \lambda\varphi$, se tiene:

$$\psi_u = \rho \cdot \varphi_u = \rho \cdot \lambda\varphi = \lambda \cdot \rho\varphi = \lambda\psi,$$

de ahí nuestra afirmación.

DEFINICIÓN X.2.1

Sea u un endomorfismo del espacio vectorial E de dimensión n . Se llama **determinante de u** , y se designa por medio de $\det(u)$, al escalar que para toda n -pla (x_i) de elementos de E , verifica la relación:

$$(\forall \varphi \in \Lambda^{*n}(E) \setminus \{0\}), \quad \varphi(u(x_1), \dots, u(x_n)) = \det(u) \varphi(x_1, \dots, x_n).$$

Si (x_1, \dots, x_n) designa una base (e_1, \dots, e_n) de E , y φ la forma n -lineal alternada Δ que verifica $\Delta(e_1, \dots, e_n) = 1$, se tiene:

$$\det(u) = \Delta(\varphi(e_1), \dots, \varphi(e_n)).$$

Inversamente, si x_1, \dots, x_n son n vectores cualesquiera de E , el determinante de estos vectores respecto de una base (e_1, \dots, e_n) de E es igual al determinante del endomorfismo u de E definido por $u(e_i) = x_i$ ($i = 1, 2, \dots, n$) (cf. Definición X.1.4).

TEOREMA X.2.1

- $$\left\| \begin{array}{l} \text{a) El determinante de la aplicación identidad es 1.} \\ \text{b) Si } v, u \in \mathcal{L}(E), \det(v \circ u) = \det(v) \det(u). \\ \text{c) El endomorfismo } u \text{ de } E \text{ es invertible si, y sólo si, } \det(u) \neq 0. \end{array} \right.$$

Demostración

a) Es evidente.

b) Sean x_1, \dots, x_n elementos de E . Por definición, para toda forma $\varphi \in \Lambda^{*n}(E)$, podemos escribir:

$$\begin{aligned} \varphi(v \circ u(x_1), \dots, v \circ u(x_n)) &= \det(v \circ u) \varphi(x_1, \dots, x_n) \quad \text{por una parte,} \\ \varphi(v \circ u(x_1), \dots, v \circ u(x_n)) &= \det(v) \varphi(u(x_1), \dots, u(x_n)) \\ &= \det(v) \cdot \det(u) \cdot \varphi(x_1, \dots, x_n) \quad \text{por otra,} \end{aligned}$$

de donde (eligiendo φ y x_1, \dots, x_n tales que $\varphi(x_1, \dots, x_n) \neq 0$):

$$\det(v \circ u) = \det(v) \cdot \det(u).$$

c) Si $u \in GL(E)$, $\det(u \circ u^{-1}) = \det(\text{id}_E) = 1 = \det(u) \cdot \det(u^{-1})$, de donde $\det(u) \neq 0$.

Por otro lado, si $u \notin GL(E)$, el rango de u es $< n$ (Teorema VIII.4.3). Cualesquiera que sean los vectores x_1, \dots, x_n de E , los vectores $u(x_1), \dots, u(x_n)$ están, pues, ligados; y (según X.1.4) se tiene:

$$0 = \varphi(u(x_1), \dots, u(x_n)) = \det(u) \varphi(x_1, \dots, x_n)$$

para toda forma n -lineal alternada φ en E , de donde $\det(u) = 0$. c.q.d.

COROLARIO

- $$\left\| \begin{array}{l} \text{Para que } n \text{ vectores } x_1, \dots, x_n \text{ de un espacio vectorial } E, \text{ de dimensión } n, \\ \text{constituyan una base de } E, \text{ es necesario y suficiente que su determinante} \\ \text{respecto de una base } (e_1, \dots, e_n) \text{ de } E, \text{ sea no nulo.} \end{array} \right.$$

En efecto, los vectores x_1, \dots, x_n constituyen una base de E si, y sólo si, el endomorfismo u definido por $u(e_i) = x_i$ ($1 \leq i \leq n$) es *invertible*.

Este resultado se puede establecer también directamente a partir de X.1.3 y X.1.4.

Determinante de una matriz

Si (e_1, \dots, e_n) es una base del espacio vectorial E , y si u designa el endomorfismo de E definido por

$$u(e_j) = \sum_{i=1}^n a_{ij} e_i \quad (1 \leq j \leq n),$$

la relación (4) del § 1 nos da inmediatamente la expresión de u por medio de la matriz $[a_{ij}]$:

$$(1) \quad \det(u) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}.$$

Esto nos lleva a establecer la siguiente definición:

DEFINICIÓN X.2.2

Sea $A = [a_{ij}]$ una matriz cuadrada de orden n . Se le llama *determinante* de A , y se designa por $\det(A)$ (o \boxed{A}), al escalar

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}.$$

Notación práctica: $\boxed{a_{ij}}$, o, en forma desarrollada,

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & & \vdots \\ a_{n,1} & \dots & \dots & a_{n,n} \end{vmatrix}.$$

El $\det(A)$ es un polinomio homogéneo de peso n respecto del conjunto de las variables (a_{ij}) , y es un polinomio (no homogéneo) de grado 1 respecto a cada una de dichas variables. El polinomio $\det(A)$ contiene exactamente $n!$ monomios (puesto que $\text{card}(\mathfrak{S}_n) = n!$). Se puede demostrar (cf. ejercicios) que el polinomio $\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) X_{\sigma(1),1} \dots X_{\sigma(n),n}$, respecto de las variables (X_{ij}) , es *irreducible*.

La fórmula (1) prueba que *el determinante de un endomorfismo de un espacio vectorial E de dimensión finita, es igual al determinante de su matriz en toda base de E .*

En particular, si $A \in M_n(K)$, y si u_A designa al endomorfismo de K^n cuya matriz, en la base canónica, es A , se tiene

$$\det(A) = \det(u_A).$$

Se tiene además:

El determinante de la matriz A es igual al determinante de los vectores columna de A respecto de la base canónica de K^n .

Teniendo en cuenta que $A \mapsto u_A$ es un isomorfismo de $M_n(K)$ en $\mathcal{L}(K^n)$, se puede aplicar el teorema X.2.1, y se obtiene:

TEOREMA X.2.2

- | | |
|--|---|
| | <p>a) $\det(I_n) = 1$.</p> <p>b) Si A y B son matrices cuadradas de orden n, se tiene:</p> $\det(AB) = \det(A) \det(B).$ <p>c) Una matriz cuadrada A de orden n es invertible si, y sólo si, $\det(A) \neq 0$.</p> |
|--|---|

Por otro lado, según las definiciones y resultados del § 1, el determinante de una matriz cuadrada A de orden n es una forma n -lineal alternada de las columnas. En virtud de las propiedades de las formas alternadas, se obtienen las siguientes propiedades (en donde las columnas de A se identifican con vectores de K^n):

1) Si se efectúa sobre las columnas de A una permutación σ , $\det A$ se transforma en $\varepsilon(\sigma) \det(A)$.

2) Si las columnas de A están ligadas, $\det(A) = 0$. Y si las columnas de A son libres, $\det(A) \neq 0$ (hallamos así, de nuevo, c) del teorema X.2.2).

3) $\det(A)$ es lineal respecto de una columna cualquiera (permaneciendo fijas las otras).

4) $\det(A)$ no cambia si añadimos a una columna de A una combinación lineal cualquiera de las *otras* columnas. En efecto, si c_1, \dots, c_n designan las columnas, se tiene:

$$\det\left(\left[c_1 + \sum_{i>1} \lambda_i c_i, c_2, \dots, c_n\right]\right) = \det([c_1, \dots, c_n]) + \det\left(\left[\sum_{i>1} \lambda_i c_i, c_2, \dots, c_n\right]\right)$$

y el último término de esta expresión es nulo puesto que las columnas de

$$\left[\sum_{i>1} \lambda_i c_i, c_2, \dots, c_n \right]$$

están ligadas.

● 5) Finalmente, para todo escalar λ , se tiene, si A es de orden n :

$$\det(\lambda A) = \lambda^n \det(A) .$$

TEOREMA X.2.3

|| *El determinante de una matriz cuadrada es igual al de su traspuesta.*

Demostración. Sea $A = [a_{ij}]$ una (n, n) -matriz:

$$(2) \quad \det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} ,$$

$$(3) \quad \det({}^t A) = \sum_{\rho \in \mathfrak{S}_n} \varepsilon(\rho) a_{1,\rho(1)} \cdots a_{n,\rho(n)} .$$

El producto $a_{1,\rho(1)} \cdots a_{n,\rho(n)}$ permanece invariante si sometemos sus factores a una permutación cualquiera. Para cada $\sigma \in \mathfrak{S}_n$, se tiene, pues:

$$a_{1,\rho(1)} \cdots a_{n,\rho(n)} = a_{\sigma(1),\rho\sigma(1)} \cdots a_{\sigma(n),\rho\sigma(n)}$$

de donde, tomando $\sigma = \rho^{-1} : a_{1,\rho(1)} \cdots a_{n,\rho(n)} = a_{\rho^{-1}(1),1} \cdots a_{\rho^{-1}(n),n} ;$

(3) se escribirá:

$$\det({}^t A) = \sum_{\rho \in \mathfrak{S}_n} \varepsilon(\rho) a_{\rho^{-1}(1),1} \cdots a_{\rho^{-1}(n),n} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} ,$$

teniendo en cuenta que $\varepsilon(\rho) = \varepsilon(\rho^{-1})$, y que $\rho \mapsto \rho^{-1}$ es una biyección de \mathfrak{S}_n en sí mismo. c.q.d.

En consecuencia, *todo lo que se ha establecido para las columnas de una matriz cuadrada (a continuación del teorema X.2.3) es válido para las filas.*

Desarrollo de un determinante

El teorema que sigue es de gran utilidad:

TEOREMA X.2.4

Sea $M = [a_{ij}]$ una matriz cuadrada de orden n , tal que $a_{ij} = 0$ para $(1 \leq p \text{ y } j < p + 1)$, por lo tanto, de la forma:

$$M = \begin{bmatrix} A & C \\ 0 \dots 0 & B \end{bmatrix},$$

en donde A es cuadrada de orden p , B es cuadrada de orden $q = n - p$, y en donde C es una (p, q) -matriz arbitraria.

En estas condiciones y con estas notaciones, se tiene:

$$\det(M) = \det(A) \det(B).$$

Demostración. Fijada la matriz C , definimos una aplicación δ de $M_p(K) \times M_q(K)$ en K por medio de la fórmula

$$\delta(X, Y) = \det(M_{X,Y}), \text{ en donde } M_{X,Y} \text{ es la matriz } \begin{bmatrix} X & C \\ 0 \dots 0 & Y \end{bmatrix}.$$

δ es una función p -lineal y alternada de las columnas de X , luego existe una función $\varphi(Y)$ que sólo depende de la matriz Y tal que:

$$\delta(X, Y) = \det(X) \cdot \varphi(Y).$$

Si hacemos $X = I_p$, se obtiene: $\varphi(Y) = \delta(I_p, Y)$. Esto demuestra que $\varphi(Y)$ es una función q -lineal alternada de las filas de Y , por lo tanto existe una constante k tal que $\varphi(Y) = k \det(Y)$. Si hacemos $Y = I_q$, obtenemos

$$k = \varphi(I_q) = \delta(I_p, I_q).$$

Recapitulando, se tiene:

$$\delta(X, Y) = \det(X) \cdot \det(Y) \delta(I_p, I_q).$$

Falta calcular entonces $\delta(I_p, I_q)$, que es el determinante de $N = \begin{bmatrix} I_p & C \\ 0 \dots 0 & I_q \end{bmatrix}.$

Pero se ve que N se obtiene a partir de I_n añadiendo a las q últimas columnas combinaciones lineales de las p primeras. De donde:

$$\det(N) = \det(I_n) = 1.$$

Razonando por recurrencia, se obtiene el

COROLARIO

Si M es una matriz de la forma:

$$\begin{bmatrix} A_1 & \text{términos} \\ 0 & A_2 \text{ cualesquiera} \\ \vdots & \vdots \\ 0 & \dots\dots\dots 0 & A_k \end{bmatrix},$$

en donde los A_i son matrices cuadradas cualesquiera, se tiene:

$$\det(M) = \det(A_1) \det(A_2) \dots \det(A_k).$$

DEFINICIÓN X.2.3

Sea $A = [a_{ij}]$ una (n, n) -matriz. Se llama **menor** relativo al término a_{ij} al determinante de la $(n-1, n-1)$ -matriz obtenida suprimiendo en A la i -ésima fila y la j -ésima columna.

TEOREMA X.2.5

(Desarrollo de un determinante según una fila o una columna.) Sea $A = [a_{ij}]$ una (n, n) -matriz. Cualesquiera que sean los índices i y j se tiene

$$(4) \quad \det(A) = \sum_{k=1}^n (-1)^{k+j} a_{kj} \Delta_{kj} = \sum_{k=1}^n (-1)^{i+k} a_{ik} \Delta_{ik}$$

en donde Δ_{ij} designa el menor relativo a a_{ij} .

Demostración. En virtud del teorema X.2.3 es suficiente demostrar la afirmación relativa a las columnas. Podemos llevar la columna j al lugar de la columna 1, sin que se modifiquen los menores de los a_{ij} , efectuando la permutación circular

$\begin{pmatrix} 1 & 2 & \dots & j \\ j & 1 & 2 & \dots & j-1 \end{pmatrix}$ con las columnas, lo que equivale a efectuar $j-1$ trasposi-

ciones (cambio de la columna j con la columna $j - 1$, luego con $j - 2$, etc.). De todo ello resulta que es suficiente probar (4) cuando $j = 1$.

Descomponiendo la primera columna en la suma de las n columnas,

$$\begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ a_{21} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ a_{n,1} \end{bmatrix},$$

se obtiene:

$$\det(A) = \sum_{k=1}^n \det(M_k),$$

en donde

$$M_k = \begin{bmatrix} 0 & a_{1,2} & \dots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{k,1} & & & \\ 0 & & & \\ \vdots & & & \\ 0 & a_{n,2} & \dots & a_{n,n} \end{bmatrix}.$$

Efectuemos sobre las k primeras filas de M_k la permutación circular $\begin{pmatrix} 1 & 2 & \dots & k \\ k & 1 & 2 & \dots & k-1 \end{pmatrix}$, de signatura $(-1)^{k-1} = (-1)^{k+1}$, y se tiene $\det(M_k) = (-1)^{k+1} \det(M'_k)$, en donde

$$M'_k = \begin{bmatrix} a_{k,1} & a_{k,2} & \dots & a_{k,n} \\ 0 & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{k-1,2} & \dots & \\ & a_{k+1,2} & \dots & \\ \vdots & \vdots & & \\ 0 & a_{n,2} & \dots & a_{n,n} \end{bmatrix}.$$

Según el teorema X.2.4, es $\det(M'_k) = a_{k,1} \Delta_{k,1}$. Se deduce fácilmente que

$$\det(A) = \sum_{k=1}^n (-1)^{k+1} a_{k,1} \Delta_{k,1} \cdot \text{c.q.d.}$$

Al término $(-1)^{i+j} \Delta_{ij}$ se le llama el **adjunto** ⁽¹⁾ de a_{ij} .

Inversa de una matriz

DEFINICIÓN X.2.4

Sea $A = [a_{ij}]$ una (n, n) -matriz ($n > 1$). La **matriz complementaria** de A es la **traspuesta** de la matriz de los adjuntos de A , es decir, la matriz ${}^t[A_{ij}]$, en donde A_{ij} es el adjunto de a_{ij} . Notación; \tilde{A} .

TEOREMA X.2.6

Para toda matriz A , cuadrada y de orden n , los productos $A \cdot \tilde{A}$ y $\tilde{A} \cdot A$ son iguales a la matriz $(\det A) I_n$.

Demostración. Hacemos $A = [a_{ij}]$, $A \tilde{A} = [c_{ij}]$ y designamos por A_{ij} al adjunto de a_{ij} . Por definición,

$$c_{ij} = \sum_{k=1}^n a_{ik} A_{jk}.$$

Para $i = j$, c_{ij} es precisamente el desarrollo de $\det(A)$ según la fila i .

Para $i \neq j$, c_{ij} es el desarrollo de un determinante que posee dos filas iguales, y, por lo tanto, vale 0.

De donde: $c_{ij} = \delta_{ij} \det(A)$. Demostración análoga para $\tilde{A} \cdot A$. c.q.d.

Con estas notaciones, el teorema X.2.6 se traduce a las relaciones

$$(5) \quad \sum_{k=1}^n a_{ik} A_{jk} = \sum_{k=1}^n A_{ki} a_{kj} = \delta_{ij} \det(A) \quad (i, j = 1, 2, \dots, n).$$

Del teorema X.2.6 resulta que, si $\det(A) \neq 0$, A es invertible, y

$$A^{-1} = \frac{1}{\det(A)} \tilde{A}.$$

Se tiene así una nueva demostración del teorema X.2.2 c).

⁽¹⁾ El autor los denomina **cofactores**. (N. del T.).

Regla práctica

Si hacemos $A^{-1} = [\alpha_{ij}]$, se tiene:

$$\alpha_{ij} = (-1)^{i+j} \frac{\Delta_{ji}}{\det(A)} = \frac{A_{ji}}{\det(A)}.$$

Se obtiene la inversa de una matriz invertible A de orden $n > 1$ formando la traspuesta \tilde{A} de la matriz de los adjuntos de A , y dividiendo todos los coeficientes de \tilde{A} por $\det(A)$.

(Para $n = 1$, la inversa de la matriz $[a]$, en donde $a \in K^*$, es evidentemente la matriz $\left[\frac{1}{a}\right]$).

Determinantes con elementos en un anillo conmutativo

Sea A un anillo conmutativo unífero (cf. Nota, p. 350).

Si E es un A -módulo que posee una base finita, la definición X.2.1 es válida para todo $u \in \mathcal{L}(E)$, y permite definir el determinante de u . En el teorema X.2.1 subsisten las conclusiones *a*) y *b*). La definición X.2.2 es la misma, y nos proporciona la noción de determinante de una matriz cuadrada con elementos en A . Las conclusiones *a*) y *b*) del teorema X.2.2 subsisten; así como los enunciados que le siguen, excepto 2). El teorema X.2.3 permanece inalterado, así como X.2.4 y X.2.5 (demostraciones idénticas).

Si extendemos la definición X.2.4 a este ámbito más general, vemos que el teorema X.2.6 permanece válido. Esto permite generalizar los teoremas X.2.1 *c*) y X.2.2 *c*) en la forma siguiente:

TEOREMA X.2.7

|| Sea M una (n, n) -matriz con elementos en un anillo conmutativo A . Para que M sea invertible, es necesario y suficiente que $\det(M)$ sea un elemento invertible de A .

Demostración. Si M es invertible, se tiene

$$MM^{-1} = I_n, \text{ de donde } \det(M) \det(M^{-1}) = \det(I_n) = 1,$$

luego $\det(M)$ es invertible en A .

Si $\det(M)$ es invertible en A , la fórmula $\tilde{M}M = M\tilde{M} = (\det(M))I_n$ prueba que M es invertible y que $M^{-1} = [(\det(M))]^{-1} \cdot \tilde{M}$. c.q.d.

Se deduce sin dificultad la formulación equivalente que sigue:

TEOREMA X.2.8

|| Sean A un anillo conmutativo unífero y E un A -módulo que admita una base finita. Para que un endomorfismo u de E sea invertible, es necesario y suficiente que $\det(u)$ sea invertible en A .

Aplicación. Estas consideraciones se aplican a las matrices con coeficientes enteros (caso en que $A = \mathbf{Z}$). Se aplican también a las matrices con coeficientes en el anillo $K[X_1, \dots, X_p]$ de los polinomios con p variables (en donde p es un entero cualquiera, y K un cuerpo conmutativo). En este caso las relaciones (5) son identidades formales de polinomios. Esta observación es la base de la demostración del teorema de Hamilton-Cayley (§ XI.3) en la cual tendremos que considerar matrices con coeficientes en el anillo $K[X]$.

En particular, podemos tomar como anillo A el anillo

$$K[(a_{ij})] \quad (i, j = 1, 2, \dots, n)$$

de los polinomios en las n^2 variables a_{ij} . Vemos así que las relaciones (5) son, de hecho, *identidades formales* (y no sólo funcionales) *entre polinomios de las variables* a_{ij} (los adjuntos A_{ij} son polinomios homogéneos de grado $n - 1$ de estas variables, y $\det(A)$ es un polinomio homogéneo de grado n).

Podríamos también haber obtenido esta última afirmación examinando atentamente las demostraciones de los teoremas X-2.4 al X.2.6.

§ X.3 EJEMPLOS DE CÁLCULO DE DETERMINANTES

1) El determinante de orden 2:

$$\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix}, \text{ es igual a } ad - bc.$$

2) Desarrollando el determinante general de orden 3:

$$\Delta = \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix},$$

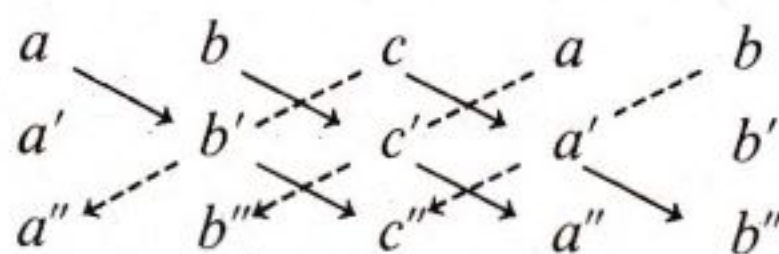
según los elementos de la primera columna, y teniendo en cuenta el ejemplo 1), se obtiene:

$$\Delta = a(b' c'' - b'' c') + a'(b'' c - b c'') + a''(b c' - b' c),$$

o sea

$$\Delta = ab' c'' + a' b'' c + a'' b c' - ab'' c' - a' b c'' - a'' b' c.$$

Para retener este último desarrollo se recurre a la regla mnemotécnica llamada «de Sarrus», ilustrada por el esquema que sigue: se repiten las dos primeras columnas de la matriz y se escriben todos los productos de tres términos situados en una misma diagonal. Los productos que corresponden a las flechas Noroeste-Sudeste deben ir precedidos del signo +, los que corresponden a las flechas Nordeste-Sudoeste deben ir precedidos del signo -.



3) Calcular el determinante de orden 3:

$$\Delta = \begin{vmatrix} (b+c)^2 & a^2 & a^2 \\ b^2 & (c+a)^2 & b^2 \\ c^2 & c^2 & (a+b)^2 \end{vmatrix}, \text{ en donde } a \text{ y } b \text{ son no nulos.}$$

Solución. Se restará la última columna a las otras dos, de donde

$$\Delta = (a+b+c)^2 \Delta',$$

con

$$\Delta' = \begin{vmatrix} b+c-a & 0 & a^2 \\ 0 & c+a-b & b^2 \\ c-a-b & c-a-b & (a+b)^2 \end{vmatrix}.$$

En Δ' , a la tercera fila se le resta la suma de la primera y la segunda, de donde

$$\begin{aligned} \Delta' &= \begin{vmatrix} b+c-a & 0 & a^2 \\ 0 & c+a-b & b^2 \\ -2b & -2a & 2ab \end{vmatrix} \\ &= \frac{1}{ab} \begin{vmatrix} a(b+c-a) & 0 & a^2 \\ 0 & b(c+a-b) & b^2 \\ -2ab & -2ab & 2ab \end{vmatrix} = \frac{1}{ab} \Delta''. \end{aligned}$$

En Δ'' , se añade la última columna a las otras dos, de donde:

$$\begin{aligned}\Delta' &= \frac{1}{ab} \begin{vmatrix} a(b+c) & a^2 & a^2 \\ b^2 & b(c+a) & b^2 \\ 0 & 0 & 2ab \end{vmatrix} = 2 \begin{vmatrix} a(b+c) & a^2 \\ b^2 & b(c+a) \end{vmatrix} \\ &= 2ab \begin{vmatrix} b+c & a \\ b & c+a \end{vmatrix} = 2abc(a+b+c).\end{aligned}$$

Finalmente, se tiene:

$$\Delta = \Delta'(a+b+c)^2, \text{ o sea } \Delta = 2abc(a+b+c)^3.$$

4) En un anillo conmutativo unífero cualquiera, sea:

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix}.$$

Desarrollando Δ según los elementos de su primera columna, después, en la expresión obtenida, desarrollando los determinantes de orden 3 según su primera columna, se obtiene Δ en función de los menores de orden 2 de su matriz. El resultado así obtenido es digno de mención, y sólo hace intervenir estos menores. Se obtiene:

$$\begin{aligned}(1) \quad \Delta &= \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \cdot \begin{vmatrix} a_3 & d_3 \\ a_4 & d_4 \end{vmatrix} + \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} \cdot \begin{vmatrix} b_3 & d_3 \\ b_4 & d_4 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_3 & d_3 \\ c_4 & d_4 \end{vmatrix} \\ &+ \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} \cdot \begin{vmatrix} b_3 & c_3 \\ b_4 & c_4 \end{vmatrix} + \begin{vmatrix} b_1 & d_1 \\ b_2 & d_2 \end{vmatrix} \cdot \begin{vmatrix} c_3 & a_3 \\ c_4 & a_4 \end{vmatrix} + \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} \cdot \begin{vmatrix} a_3 & b_3 \\ a_4 & b_4 \end{vmatrix}.\end{aligned}$$

Es posible retener este resultado por medio de la siguiente definición: para cada determinante δ de orden 2 de la matriz de Δ , llamamos *adjunto* de este menor al menor de orden 2 formado con los elementos que no están ni en las filas ni en las columnas de δ . Δ es entonces la suma de seis productos:

1.º Los productos por sus adjuntos de los menores de orden 2 formados con $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix}$ suprimiendo sucesivamente las columnas a, b, c ; el segundo menor va afectado del signo $-$.

2.º Los productos por sus adjuntos de los menores de orden 2 formados con $\begin{bmatrix} a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{bmatrix}$ suprimiendo sucesivamente las columnas a, b, c ; el segundo menor va afectado del signo $-$.

En la expresión (1) hacemos $a_1 = a_3, a_2 = a_4, b_1 = b_3, b_2 = b_4, \dots, d_1 = d_3, d_2 = d_4$. Se obtiene el desarrollo de un determinante nulo (puesto que Δ tiene entonces dos pares de filas iguales). De donde se obtiene la relación (dividiendo por 2, si ello está permitido en el anillo de base):

$$(2) \quad \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \cdot \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} + \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} \cdot \begin{vmatrix} b_1 & d_1 \\ b_2 & d_2 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} = 0,$$

que prueba que los menores de orden 2 de la matriz $\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \end{bmatrix}$ no son algebraicamente independientes. La relación (2) permite definir las coordenadas plückerianas de una recta en Geometría proyectiva.

5) *Determinante de Vandermonde*. Si A es un anillo conmutativo unífero, sea:

$$\Delta_n = \det(M_n), \text{ en donde } M_n = \begin{bmatrix} 1 & x_1 & (x_1)^2 & \dots & (x_1)^{n-1} \\ 1 & x_2 & & & \vdots \\ \vdots & & & & \vdots \\ 1 & x_n & \dots & \dots & (x_n)^{n-1} \end{bmatrix} \quad \begin{matrix} x_i \in A \\ (1 \leq i \leq n) \end{matrix}.$$

Vamos a establecer la relación:

$$(3) \quad \Delta_n = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

El resultado es evidente para $n = 2$, y razonamos por recurrencia sobre n . Suponemos (3) verdadero, y calculamos $\Delta_{n+1} = \det(M_{n+1})$:

$$\Delta_{n+1} = \begin{vmatrix} 1 & x_1 & \dots & (x_1)^n \\ \vdots & & & \vdots \\ 1 & x_{n+1} & \dots & (x_{n+1})^n \end{vmatrix}$$

Restamos la última fila de todas las demás,

$$\Delta_{n+1} = \begin{vmatrix} 0 & x_1 - x_{n+1} & x_1^2 - x_{n+1}^2 & \dots & x_1^n - x_{n+1}^n \\ \vdots & & & & \vdots \\ 0 & x_n - x_{n+1} & \dots & \dots & (x_n)^n - (x_{n+1})^n \\ 1 & x_{n+1} & \dots & \dots & (x_{n+1})^n \end{vmatrix}.$$

Desarrollamos según la primera columna y ponemos de manifiesto el factor $x_1 - x_{n+1}$ en la fila i ($1 \leq i \leq n$), y se obtiene:

$$(4) \quad \Delta_{n+1} = \prod_{i=1}^n (x_{n+1} - x_i) \cdot D,$$

con

$$D = \begin{vmatrix} 1 & x_1 + x_{n+1} & \dots & x_1^{n-1} + x_1^{n-2} x_{n+1} + \dots + (x_{n+1})^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n + x_{n+1} & \dots & \dots \end{vmatrix}.$$

En el determinante D , restamos de la última columna, la penúltima multiplicada por x_{n+1} , luego a la $(n-1)$ -columna, la $(n-2)$ -ésima multiplicada por x_{n+1} , etc., se obtiene:

$$D = \begin{vmatrix} 1 & x_1 & (x_1)^2 & \dots & (x_1)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & \dots & (x_n)^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i) \quad (\text{según (3)});$$

de donde, utilizando (4):

$$\Delta_{n+1} = \prod_{1 \leq i < j \leq n+1} (x_j - x_i),$$

lo cual demuestra que la relación es verdadera para el orden $n+1$. c.q.d.

Otro método. Consideremos Δ_n como un polinomio respecto de las variables $(x_i)_{1 \leq i \leq n}$. Si damos a x_i y a x_j ($i \neq j$) valores iguales, Δ_n se anula (2 filas iguales) luego $x_i - x_j$ ($i \neq j$) divide a Δ_n (cf. p. 158). Puesto que los $x_i - x_j$ son primos entre sí, Δ_n es divisible por el producto de los $x_i - x_j$ (§ XIV.2). Por razones de grado, se tiene

$$\Delta_n = C \prod_{i < j} (x_j - x_i),$$

en donde C es una constante; comparando, por ejemplo, los términos en $(x_1)^{n-1}$ vemos que $C = 1$. (Este método presupone que A es un cuerpo.)

6) En un cuerpo conmutativo, calculamos el determinante de orden n :

$$\Delta = \begin{vmatrix} r_1 & a & \dots & a \\ b & r_2 & & \vdots \\ \vdots & & & a \\ b & \dots & b & r_n \end{vmatrix} \quad (a \neq b),$$

Hacemos

$$\Delta(x) = \begin{vmatrix} r_1 + x & a + x & \dots & a + x \\ b + x & r_2 + x & & \\ \vdots & & & \\ b + x & \dots & \dots & r_n + x \end{vmatrix}.$$

Consideremos la k -ésima columna como la suma de las dos columnas

$$\begin{bmatrix} a \\ \vdots \\ r_k \\ b \\ \vdots \\ b \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} x \\ \vdots \\ x \end{bmatrix},$$

y apliquemos la multilinealidad: $\Delta(x)$ es suma de 2^n determinantes. Entre todos estos determinantes, los formados con dos columnas de la forma $\begin{bmatrix} x \\ \vdots \\ x \end{bmatrix}$ son nulos.

En otras palabras, $\Delta(x)$ es un polinomio de *primer grado* en x ; hacemos:

$$\Delta(x) = Ax + B.$$

Para calcular las constantes A y B , reemplazamos sucesivamente x por $-a$ y por $-b$. Poniendo $\varphi(x) = (r_1 - x) \dots (r_n - x)$, se obtiene:

$$-aA + B = \varphi(a), \quad -bA + B = \varphi(b), \quad \text{de donde} \quad A = B = \frac{a\varphi(b) - b\varphi(a)}{a - b}.$$

Cuando $a = b$, el cálculo directo de Δ es más fácil, y lo dejamos al lector.

Sin embargo, del resultado hallado para $a \neq b$, vamos a deducir el resultado para $a = b$. Cuando el cuerpo no tiene característica 2, se puede escribir

$$\varphi(b) = \varphi(a) + (b - a) \varphi'(a) + (b - a)^2 \psi(a),$$

en donde ψ es cierto polinomio.

De donde, para $b \neq a$:

$$(5) \quad \Delta = \varphi(a) - a\varphi'(a) + (a - b) \psi(a).$$

Fijemos $a \in K$, supongamos que el cuerpo es infinito, y consideremos los dos miembros de (5) como polinomios en b . Estos polinomios son iguales para una infinidad de valores, puesto que lo son para $b \neq a$. Son, pues, formalmente iguales. Por lo tanto, toman el mismo valor para $b = a$; el valor de Δ para $b = a$ es, pues, $\varphi(a) - a\varphi'(a)$.

7) Determinante de la *matriz circulante* con elementos en el cuerpo \mathbf{C} :

$$M = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \vdots & & & \vdots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix}.$$

Cada fila se deduce de la precedente por una permutación circular.

Sumando las columnas vemos que $a_0 + a_1 + \dots + a_{n-1}$ divide a $\Delta = \det(M)$ (considerado como polinomio de a_0, a_1, \dots, a_{n-1}). En general, si ζ es una raíz n -ésima de 1, vemos fácilmente que Δ es divisible por $A_i = \sum_{0 \leq i \leq n-1} \zeta^i a_i$. Podemos concluir entonces que

$$\Delta = C \prod_{i=1}^n A_i,$$

en donde C es una constante, gracias a la factorización de los polinomios con varias variables (y teniendo en cuenta el grado de M), (cf. Cap. XIV).

Vamos a encontrar de nuevo este resultado sin necesidad de recurrir a la teoría de los polinomios. Sea ω una raíz primitiva n -ésima de 1 (por ejemplo $\omega = e^{2\pi i/n}$), y sea $\Omega = [\omega^{(i-1)(j-1)}]$ (matriz de Vandermonde de las raíces n -ésimas de 1). Las raíces n -ésimas $(\omega^{i-1})_{1 \leq i \leq n}$ son distintas y, por lo tanto, se tiene $\det(\Omega) \neq 0$. Vamos a calcular $M\Omega$, y $\det(M\Omega)$. De lo que deduciremos $\det(M)$, gracias a la fórmula:

$$(5) \quad \det(M\Omega) = \det(M) \det(\Omega).$$

Designamos por \overline{m} la clase del entero m mód (n) . La matriz M se escribe:

$$M = [a_{\overline{j-i}}]_{1 \leq i \leq n, 1 \leq j \leq n}.$$

Un cálculo directo proporciona: $M\Omega = [b_{ij}]_{1 \leq i \leq n, 1 \leq j \leq n}$, con,

$$(6) \quad b_{ij} = \sum_{k=1}^n a_{\overline{k-i}} \omega^{(k-1)(j-1)} = \omega^{(i-1)(j-1)} \sum_{k=1}^n a_{\overline{k-i}} \omega^{(j-1)(k-i)} = \omega^{(i-1)(j-1)} C_{ij},$$

en donde $C_{ij} = \sum_{k=1}^n a_{k-i} \omega^{(j-1)(k-i)}$. Pero, observando que para todo entero m y toda raíz n -ésima de la unidad ζ , el número ζ^m depende únicamente de \bar{m} , vemos que:

$$C_{ij} = \sum_{\lambda=0}^{n-1} a_{\lambda} \omega^{\lambda(j-1)}, \quad \text{y } C_{ij} \text{ sólo depende de } j;$$

Indicaremos por C_j el valor común de los C_{ij} ($1 \leq i \leq n$).

Para calcular $\det(M\Omega) = \det([b_{ij}])$, podemos, según (6), introducir el factor C_j en la j -ésima columna. Obtenemos:

$$\det(M\Omega) = \left(\prod_{j=1}^n C_j \right) \det(\Omega).$$

Comparando con (5), y teniendo presente que $\det(\Omega) \neq 0$, se deduce:

$$\det(M) = \prod_{j=1}^n C_j,$$

o sea, intercambiando las notaciones:

$$\boxed{\det(M) = \prod_{\mu=0}^{n-1} \left(\sum_{\lambda=0}^{n-1} \omega^{\lambda\mu} a_{\lambda} \right)}.$$

Nota. El método anterior evita el cálculo de $\det(\Omega)$. El cálculo de $(\det(\Omega))^2$ es interesante por sí mismo. El lector comprobará las fórmulas:

$$\Omega^2 = \begin{bmatrix} n & 0 & \dots & 0 \\ 0 & \dots & 0 & n \\ \vdots & & \ddots & \vdots \\ 0 & n & 0 & \dots & 0 \end{bmatrix}, \quad \text{y} \quad \det(\Omega^2) = (-1)^{(n-1)(n-2)/2} n^n.$$

§ X.4 APLICACIÓN DE LOS DETERMINANTES AL ESTUDIO DEL RANGO DE UNA MATRIZ

- Consideraremos una matriz $M = [a_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ con elementos en el **cuerpo**

conmutativo K .

Designaremos «vectores fila» (resp. «vectores columna») a las filas (resp. columnas) de la matriz M , identificadas con elementos de K^p (resp. K^n).

DEFINICIÓN X.4.1

Sea M una (n, p) -matriz. Se llama **rango** de M al rango del sistema de vectores columna de M , es decir, a la dimensión del espacio vectorial engendrado por estos vectores (cf. Definición VIII.3.2).

Cada (n, p) -matriz M se puede considerar como la matriz de una aplicación lineal φ , de K^p en K^n , en las bases canónicas de estos espacios (cap. IX, § 2). Los vectores columna de M son entonces las imágenes por φ de los vectores de la base canónica de K^p , y el rango de M es igual al rango de φ (Definición VIII. 4.1).

En general, si E, F son dos espacios vectoriales de dimensión finita, y si φ es una aplicación lineal de E en F , el rango de φ es igual al rango de su matriz M en bases cualesquiera de E, F .

Hemos establecido (Teorema VIII.5.8) que el rango de una aplicación lineal es igual al de su traspuesta, y que en las bases duales de las elegidas en E, F , la matriz de ${}^t\varphi$ es tM (Teorema IX.2.3). Se tiene, pues, el siguiente teorema:

TEOREMA X.4.1

El rango de una matriz es igual al de su traspuesta. En otras palabras, es el rango del sistema de los vectores fila de la matriz.

El rango de la matriz $[a_{ij}]_{1 \leq i \leq n, 1 \leq j \leq p}$ es, pues, igual al rango de la familia constituida por las n **formas lineales**

$$y_i = \sum_{j=1}^p a_{ij} x_j \quad (1 \leq i \leq n).$$

(Consideramos estas formas como elementos del dual de K^p , referido a la base (x_1, \dots, x_p) .)

Matriz extraída ⁽¹⁾

DEFINICIÓN X.4.2

Sea $M = [a_{ij}]$ una (n, p) -matriz, y sea
 $I = \{i_1, i_2, \dots, i_m\}$ una parte de \mathbf{N}_n^* , $i_1 < i_2 < \dots < i_m$;
 $J = \{j_1, j_2, \dots, j_q\}$ una parte de \mathbf{N}_p^* , $j_1 < j_2 < \dots < j_q$.

⁽¹⁾ También llamadas **submatrices**. (N. del T.)

La matriz **extraída** de la M , relativa a I y a J , es la matriz M_{IJ} con m filas y q columnas, de término general

$$\alpha_{\lambda\mu} = a_{i_\lambda, j_\mu} \quad (1 \leq \lambda \leq m, 1 \leq \mu \leq q).$$

En la práctica diremos que M_{IJ} se ha obtenido por medio de las filas de índice en I , y de las columnas de índice en J ; o también, que M_{IJ} se obtiene suprimiendo en M las filas cuyo índice no pertenece a I , y las columnas cuyo índice no pertenece a J .

El número de (m, q) -submatrices de M es evidentemente $\binom{n}{m} \cdot \binom{p}{q}$.

Por abuso de lenguaje, se llamará «subdeterminante de orden p » al determinante de una matriz cuadrada de orden p extraída de M .

Los subdeterminantes permiten estudiar el rango de una matriz.

Empezaremos por establecer un lema, cuyo recíproco resultará de X.4.4.

X.4.2 Sea $M = [a_{ij}]$ una (n, p) -matriz. Si existe un subdeterminante de M , de orden r , Δ_r , tal que $\Delta_r \neq 0$, el rango de M es $\geq r$.

Demostración. Podemos suponer que los vectores columna que corresponden a Δ_r , son los r primeros, y que los vectores fila que corresponden a Δ_r son los r primeros (puesto que las permutaciones de las filas y de las columnas no afectan al rango de M).

Sean v_1, \dots, v_p los vectores columna de M , y v'_1, \dots, v'_p los vectores columna «truncados» a partir del índice $r + 1$. En otras palabras, v'_i es un vector de K^r cuyas componentes son las r primeras componentes de v_i . De forma general, la aplicación $v \mapsto v'$ es lineal, epiyectiva de K^n en K^r , y su núcleo es isomorfo a K^{n-r} (es la proyección de K^n en el subespacio $K^r \times \{0\}$, paralelamente al subespacio $\{0\} \times K^{n-r}$). El determinante de v'_1, \dots, v'_r es Δ_r . Luego v'_1, \dots, v'_r son linealmente independientes en K^r (Corolario del teorema X.2.1). A fortiori v_1, \dots, v_r son linealmente independientes en K^n . c.q.d.

Sea Δ un subdeterminante de orden r de la matriz M . Decimos que un subdeterminante de M , de orden $r + 1$, es un *orlado* de Δ , si admite como subdeterminante a Δ . Se tiene la propiedad que sigue:

X.4.3 Sea $M = [a_{ij}]$ una (n, p) -matriz, y sea Δ_r un subdeterminante de M , de orden r , tal que $\Delta_r \neq 0$. Si el rango de M es $\geq r + 1$, existe un determinante, no nulo, que es orlado del Δ_r .

Demostración. Se puede suponer que Δ_r está formado con las filas y las columnas de índices $1, 2, \dots, r$, de M . Sean v_1, \dots, v_p los vectores columna de M .

Puesto que $\Delta_r \neq 0$, los vectores v_1, \dots, v_r son linealmente independientes. Puesto que el rango de M es $\geq r + 1$, existe un índice $k \geq r + 1$ tal que los vectores v_1, \dots, v_r, v_k son linealmente independientes. Sea N la matriz de columnas v_1, \dots, v_r, v_k . Es claro que N es de rango $r + 1$. Luego el sistema de los vectores fila de N es de rango $r + 1$ (cf. X.4.1).

Designemos por w_1, \dots, w_n las filas de N . Dado que $\Delta_r \neq 0$, los vectores w_1, \dots, w_r son linealmente independientes; existe, pues, un índice $l > r$ tal que el sistema (w_1, \dots, w_r, w_l) es libre.

Por lo tanto, el determinante de w_1, \dots, w_r, w_l es un subdeterminante de M de orden $r + 1$, que es orlado del Δ_r , y que es no nulo ya que (w_1, \dots, w_r, w_l) es libre.

Podemos resumir X.4.1 y X.4.2 en el teorema fundamental que sigue:

TEOREMA X.4.4

|| *Para que una matriz M sea de rango r , es necesario y suficiente*
 1) *que exista un subdeterminante Δ_r de M , de orden r , tal que $\Delta_r \neq 0$;*
 2) *que todos los orlados del Δ_r sean nulos.*

La condición 2) se supondrá verificada si r es igual al número de filas (o de columnas) de la matriz M — pues entonces no existe ningún orlado de orden $r + 1$.

COROLARIO

|| *El rango de una matriz es el orden máximo de un subdeterminante no nulo de dicha matriz.*

Nota. Supongamos conocido un determinante no nulo Δ_r , subdeterminante de orden r de la (n, p) -matriz M . Para comprobar que M es de rango r , basta asegurarse de que los orlados de Δ_r (en número de $(p - r)(n - r)$) son nulos. En particular, si estos orlados son nulos, es *inútil* estudiar los subdeterminantes de orden $> r + 1$.

Rango de un sistema de vectores

El estudio del rango de una matriz permite determinar el rango de un sistema de p vectores de un espacio vectorial E de dimensión n . En efecto, sean V_1, \dots, V_p dichos vectores: si $(a_{ij})_{1 \leq j \leq n}$ son las coordenadas de V_i en una base prefijada de E , el rango de la matriz $M = [a_{ij}]_{1 \leq i \leq p, 1 \leq j \leq n}$ es igual al rango del sistema de vectores $(V_i)_{1 \leq i \leq p}$.

§ X.5 ECUACIONES LINEALES

Un *sistema de ecuaciones lineales* consiste en el siguiente problema: dada una (m, n) -matriz $[a_{ij}]$ y los m escalares $(b_i)_{1 \leq i \leq m}$, hallar todos los sistemas $(x_j)_{1 \leq j \leq n}$ de n escalares que verifican las relaciones

$$(1) \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad (1 \leq i \leq m).$$

Por motivos de brevedad, hablaremos del *sistema de ecuaciones lineales* (1). Se llama *solución* del sistema a toda n -pla (x_1, \dots, x_n) que verifique (1).

La matriz $M = [a_{ij}]$ es la *matriz del sistema*, el rango de M es el *rango del sistema*, el vector columna $(b_i)_{1 \leq i \leq m}$ es el *segundo miembro* del sistema, las x_j son las *incógnitas*.

Si el segundo miembro es nulo, al sistema se le llama *homogéneo*.

Al sistema que se obtiene, a partir de un sistema de ecuaciones dado, reemplazando el segundo miembro por el vector columna nulo, se le llama *sistema homogéneo asociado* (al sistema dado).

El sistema (1) se puede escribir en las dos formas equivalentes que siguen:

$$\text{Forma matricial. Haciendo } M = [a_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \text{ y } \mathcal{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \mathcal{B} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

(1) equivale a:

$$(2) \quad M \mathcal{X} = \mathcal{B}, \text{ en donde } \mathcal{X} \text{ es la incógnita.}$$

Forma vectorial. Sean V_1, \dots, V_n los vectores columna de M , \mathcal{B} el segundo miembro, (1) equivale a:

$$(3) \quad x_1 V_1 + x_2 V_2 + \dots + x_n V_n = \mathcal{B} \quad (x_1, \dots, x_n, \text{ incógnitas}).$$

DEFINICIÓN X.5.1

$\left\{ \begin{array}{l} \text{Un sistema de ecuaciones lineales es } \mathbf{compatible} \text{ si admite una solución,} \\ \text{por lo menos, e } \mathbf{incompatible} \text{ en caso contrario.} \end{array} \right.$

El caso fundamental de sistema compatible es el sistema de Cramer:

DEFINICIÓN X.5.2

$\left\{ \begin{array}{l} \text{A un sistema de ecuaciones lineales se le llama de } \mathbf{Cramer} \text{ si su matriz} \\ \text{es cuadrada e invertible.} \end{array} \right.$

Un sistema de Cramer, escrito en la forma (2), se presenta en la forma:

$$M \mathcal{X} = \mathcal{B}, \quad M \text{ cuadrada invertible};$$

si multiplicamos esta relación por M^{-1} por la izquierda, obtendremos la relación equivalente (cf. asociatividad del producto de matrices)

$$(4) \quad \mathcal{X} = M^{-1} \mathcal{B}, \quad \text{de donde:}$$

TEOREMA X.5.1

\parallel *Un sistema de Cramer admite una solución única.*

En particular un sistema de Cramer homogéneo, que admita la solución $x_1 = x_2 = \dots = x_n = 0$, sólo admite esta solución: se dice que admite únicamente *la solución trivial*.

La fórmula (4) permite en teoría, y a veces en la práctica, calcular la solución de un sistema de Cramer. La escritura vectorial de un sistema lineal permite expresar la solución de un sistema de Cramer en otra forma, muchas veces más cómoda que (4).

Designemos por V_1, \dots, V_n los vectores columna de la matriz de nuestro sistema de Cramer, por \mathcal{B} el segundo miembro. Sea (x_1, \dots, x_n) la solución de este sistema. Entonces se tiene:

$$(5) \quad b = \sum_{i=1}^n x_i V_i.$$

Calculemos, teniendo en cuenta (5) y la multilinealidad:

$$\det(V_1, \dots, V_{k-1}, b, V_{k+1}, \dots, V_n) = \sum_i x_i \det(V_1, \dots, V_{k-1}, V_i, V_{k+1}, \dots, V_n);$$

el único determinante no nulo de esta suma es:

$$\det(V_1, V_2, \dots, V_{k-1}, V_k, V_{k+1}, \dots, V_n) = \Delta,$$

puesto que Δ es el determinante de la matriz del sistema. Se obtiene pues

$$(6) \quad x_k = \frac{1}{\Delta} \det(V_1, \dots, V_{k-1}, b, V_{k+1}, \dots, V_n) \quad (\text{fórmulas de Cramer}).$$

Sistema homogéneo

Consideremos el sistema homogéneo:

$$(7) \quad \sum_{j=1}^n a_{ij} x_j = 0 \quad (1 \leq i \leq m), \quad \text{y sea } r \text{ su rango.}$$

(7) es siempre compatible (puesto que admite la solución trivial).

Interpretemos $M = [a_{ij}]$ como la matriz de una aplicación lineal: $\varphi : K^n \mapsto K^m$. La resolución de (7) equivale a la búsqueda del núcleo de φ . Ahora bien, φ es de rango r . Según el corolario de VIII.3.10, se deduce:

TEOREMA X.5.2

|| El conjunto de las soluciones del sistema homogéneo (7) forma un subespacio de K^n , de dimensión $n - r$, en donde r designa el rango del sistema.

Cálculo de las soluciones

Sea Δ_r un subdeterminante de $M = [a_{ij}]$, de orden r , no nulo. A tales determinantes Δ_r se les llama *menores principales* de (7). A las ecuaciones cuyos índices son los de las filas de Δ_r , se les llama entonces *ecuaciones principales*, las incógnitas cuyos índices son los de las columnas de Δ_r son las *incógnitas principales*. En adelante podemos suponer

$$\Delta_r = \det [a_{ij}]_{\substack{1 \leq i \leq r \\ 1 \leq j \leq r}}.$$

Designemos por \mathcal{S} al subespacio de las soluciones de (7) y por \mathcal{S}_r al subespacio de las soluciones de

$$(8) \quad \sum_{j=1}^n a_{ij} x_j = 0 \quad (1 \leq i \leq r)$$

(sistema de las ecuaciones principales).

(8) es de rango r , luego $\dim(\mathcal{S}_r) = n - r$. Puesto que $\dim(\mathcal{S}) = n - r$, y que evidentemente se verifica $\mathcal{S} \subset \mathcal{S}_r$, se deduce $\mathcal{S} = \mathcal{S}_r$.

En otras palabras, la resolución de (7) equivale a la de (8).

Para resolver (8), damos valores arbitrarios a las incógnitas no principales x_{i+1}, \dots, x_n : las incógnitas principales quedan entonces determinadas por un sistema de Cramer. En resumen, hemos demostrado el:

TEOREMA X.5.3

La resolución de un sistema homogéneo equivale a la de un subsistema cualquiera de ecuaciones principales. Un sistema de ecuaciones principales se resuelve dando valores arbitrarios a las incógnitas no principales, y calculando entonces las incógnitas principales con la ayuda del sistema de Cramer así obtenido.

Caso particularmente importante de sistema homogéneo:

Sistema de n ecuaciones con $n + 1$ incógnitas, y de rango n . Tal sistema se escribe:

$$(9) \quad \sum_{j=1}^{n+1} a_{ij} x_j = 0 \quad (1 \leq i \leq n).$$

Podemos suponer entonces que $\Delta = \det([a_{ij}])$ ($i, j = 1, 2, \dots, n$) es no nulo.

Según el teorema X.4.2, las soluciones de (9) forman un subespacio de dimensión 1 de K^{n+1} ; en otras palabras, si $(\alpha_1, \dots, \alpha_{n+1})$ es una solución no trivial de (9), todas las soluciones de (9) son de la forma

$$(\lambda \alpha_1, \dots, \lambda \alpha_{n+1}), \quad \lambda \in K.$$

Por lo tanto, dicha solución se puede determinar de la forma siguiente: sean u_1, \dots, u_{n+1} términos cualesquiera, y consideremos la matriz:

$$N = \begin{bmatrix} a_{11} & \dots & a_{1,n} & a_{1,n+1} \\ \vdots & & & \vdots \\ a_{n,1} & \dots & a_{n,n} & a_{n,n+1} \\ u_1 & \dots & u_n & u_{n+1} \end{bmatrix}.$$

Sean A_1, \dots, A_{n+1} los adjuntos de u_1, \dots, u_{n+1} en esta matriz. Vamos a demostrar que (A_1, \dots, A_{n+1}) es solución de (9): En efecto, la relación

$$\sum_{j=1}^{n+1} a_{ij} A_j = 0$$

($1 \leq i \leq n$) se verifica, puesto que el primer miembro es el desarrollo del determinante de N según la fila de índice $n+1$, si reemplazamos u_1, \dots, u_{n+1} por $a_{i,1}, \dots, a_{i,n+1}$; y este determinante es nulo ya que posee dos filas iguales. Por otro lado, $A_{n+1} = \Delta \neq 0$ por hipótesis. En conclusión: *las soluciones de un sistema de la forma (9) son, con las anteriores notaciones, de la forma*

$$(\lambda A_1, \lambda A_2, \dots, \lambda A_{n+1}), \quad \lambda \in K.$$

X.5.4 Sea

$$(H) \quad \sum_{j=1}^{n+1} a_{ij} x_j = 0$$

un sistema homogéneo de rango n , con $n+1$ incógnitas; y para cada $j = 1, 2, \dots, n+1$, designemos por Δ_j el determinante de la matriz de orden n obtenida suprimiendo la columna de índice j en la matriz $[a_{ij}]$. Entonces las soluciones de (H) son de la forma $x_j = (-1)_j \lambda \Delta_j$, en donde $\lambda \in K$ es arbitrario.

Resolución de un sistema lineal: caso general

Consideremos nuevamente el sistema (1):

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad (1 \leq i \leq m), \quad \text{y sea } r \text{ su rango.}$$

Se define, de la misma manera que para los sistemas homogéneos, las nociones de determinante principal, de ecuaciones principales y de incógnitas principales.

Supondremos que el determinante principal es $\Delta_r = \det([a_{ij}])_{(i,j=1,2,\dots,r)}$. Sea M la matriz del sistema y sea b el segundo miembro. El sistema es compatible si, y sólo si, b pertenece al subespacio de K^m engendrado por los vectores columna de M ; en otras palabras, si las matrices:

$$M = [a_{ij}], \quad \text{y} \quad N = \begin{bmatrix} a_{11} & \dots & a_{1,n} & b_1 \\ \vdots & & & \vdots \\ a_{m,1} & \dots & a_{m,n} & b_m \end{bmatrix}$$

tienen el mismo rango.

Para ello es necesario y suficiente (cf. teoría de subdeterminantes orlados), que los $m-r$ orlados de Δ_r formados con la columna b sean nulos. De donde:

TEOREMA X.5.5

Sea $\sum_{j=1}^n a_{ij} x_j = b_i$ ($1 \leq i \leq m$) un sistema lineal de rango r , y supongamos que $\Delta_r = \det([a_{ij}]_{(i,j=1,2,\dots,r)})$ es un determinante principal. Para que el sistema sea compatible es necesario y suficiente que se verifiquen las $m - r$ relaciones siguientes:

$$(10) \det \begin{bmatrix} a_{1,1} & \dots & a_{1,r} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r,1} & \dots & a_{r,r} & b_r \\ a_{k,1} & \dots & a_{k,r} & b_k \end{bmatrix} = 0 \quad (r+1 \leq k \leq m).$$

A los determinantes que figuran en el primer miembro de las relaciones (10) se le llama **determinantes característicos** del sistema.

Cálculo de las soluciones de un sistema compatible

Consideremos un sistema compatible de rango r :

$$(11) \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad (1 \leq i \leq m),$$

tal que $\Delta_r \neq 0$, con $\Delta_r = \det([a_{ij}]_{1 \leq i,j \leq r})$.

Sea $\mathcal{X}_0 = (x_1^0, \dots, x_n^0)$ una solución particular de este sistema, y sea $\mathcal{X} = (x_1, \dots, x_n)$ una solución cualquiera. Es inmediato que

$$\mathcal{X} - \mathcal{X}_0 = (x_1 - x_1^0, \dots, x_n - x_n^0)$$

es una solución del sistema homogéneo asociado

$$\sum_{j=1}^n a_{ij} x_j = 0 \quad (1 \leq i \leq m), \quad \text{y recíprocamente.}$$

Es decir: Las soluciones de (11) forman un subespacio afin φ de K^n , cuya dirección es el subespacio de las soluciones del sistema homogéneo asociado. En particular, se tiene: $\dim(\mathcal{S}) = n - r$.

El sistema de las r ecuaciones principales

$$(12) \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad (1 \leq i \leq r)$$

es también de rango r . Sea \mathcal{S}_r el subespacio afín de las soluciones de (12). De las relaciones, $\dim(\mathcal{S}_r) = n - r$, y $\mathcal{S} \subset \mathcal{S}_r$, se deduce $\mathcal{S} = \mathcal{S}_r$; con otras palabras:

La resolución de (11) equivale a la de un subsistema cualquiera de ecuaciones principales.

Para resolver el sistema de las ecuaciones principales, se dan valores arbitrarios a las incógnitas no principales, y las incógnitas principales vienen dadas entonces por un sistema de Cramer, puesto que $\Delta_r \neq 0$.

Todos los resultados demostrados anteriormente se pueden condensar en el siguiente teorema:

TEOREMA X.5.6 (ROUCHÉ-FONTENÉ)

|| *Si un sistema lineal es compatible, su solución equivale a la de un sistema de ecuaciones principales cualquiera.*
 || *Para resolver un sistema de ecuaciones principales se dan valores arbitrarios a las incógnitas no principales, y las incógnitas principales se determinan entonces por medio de un sistema de Cramer.*

Ejemplos

1) Vamos a discutir el sistema general de 3 ecuaciones con 3 incógnitas x, y, z :

$$\begin{aligned} (13) \quad & ux + vy + wz = a \\ (14) \quad & u'x + v'y + w'z = a' \\ (15) \quad & u''x + v''y + w''z = a'' \end{aligned}$$

interpretando estas ecuaciones como si fuesen las ecuaciones de 3 planos afines

de K^3 . Suponemos que los vectores fila de la matriz $M = \begin{bmatrix} u & v & w \\ u' & v' & w' \\ u'' & v'' & w'' \end{bmatrix}$ son no nulos.

Si $\det(M) \neq 0$, el sistema es de Cramer, y los tres planos (13), (14) y (15) tienen un punto en común y solo uno.

Si el rango de M es igual a 2, una de las filas de M es combinación lineal de las otras dos. Podemos suponer, por ejemplo, que las filas

$$(u, v, w) \quad \text{y} \quad (u', v', w')$$

son linealmente independientes, y que (u'', v'', w'') es una combinación lineal de estas filas. En este caso, el sistema [(13), (14)] tiene una infinidad de soluciones, repartidas sobre una recta de K^3 . Por lo tanto, si el sistema es compatible, los tres planos poseen exactamente una recta de puntos comunes; y si el sistema es incompatible, dos de los tres planos son paralelos, y cada uno de ellos corta al tercero a lo largo de una recta.

Si el rango de M es igual a 1, las direcciones de los tres planos (13), (14) y (15) son idénticas, ya que las filas de M son proporcionales. Si el sistema es compatible, los tres planos se confunden. Si no, hay dos casos posibles: o bien los tres planos son distintos dos a dos, o bien exactamente dos de los tres planos se hallan confundidos.

2) Resolver y discutir, según los valores de λ , el sistema:

$$\begin{aligned} 2(\lambda + 1)x & \quad \quad \quad + 3y + \lambda z = \lambda + 4 \\ (4\lambda - 1)x + (\lambda + 1)y + (2\lambda - 1)z &= 2\lambda + 4 \\ (5\lambda - 4)x + (\lambda + 1)y + (3\lambda - 4)z &= \lambda - 1 \end{aligned}$$

El determinante del sistema es $\Delta(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3)$.

Si $(\lambda - 1)(\lambda - 2)(\lambda - 3) \neq 0$, las fórmulas de Cramer dan

$$\begin{aligned} x &= \frac{1}{\Delta(\lambda)} (2\lambda^3 - 4\lambda^2 - 27\lambda + 39), & y &= \frac{1}{\Delta(\lambda)} (3\lambda^2 + 21\lambda - 34), \\ z &= \frac{1}{\Delta(\lambda)} (-3\lambda^3 + 2\lambda^2 + 40\lambda - 49). \end{aligned}$$

Para $\lambda = 1$, el sistema se escribe:

$$\begin{aligned} 4x + 3y + z &= 5 \\ 3x + 2y + z &= 6 \\ x + 2y - z &= 0. \end{aligned}$$

El rango es 2. Añadiendo la tercera ecuación a cada una de las restantes, se obtiene:

$$5x + 5y = 5, \quad 4x + 4y = 6$$

y el sistema es incompatible.

Para $\lambda = 2$, el sistema se escribe

$$\begin{aligned} 6x + 3y + 2z &= 6 \\ 7x + 3y + 3z &= 8 \\ 6x + 3y + 2z &= 1, \end{aligned}$$

que es evidentemente incompatible.

Para $\lambda = 3$, el sistema se escribe:

$$\begin{aligned} 8x + 3y + 3z &= 7 \\ 11x + 4y + 5z &= 7 \\ 11x + 4y + 5z &= 2, \end{aligned}$$

y también es incompatible.

3) A veces es posible evitar el uso —pesado— de las fórmulas de Cramer. Consideremos el sistema siguiente, en que a_1, \dots, a_n son no nulos, en el supuesto de que constituya un sistema de Cramer:

$$\begin{aligned} (1 + a_1)x_1 + x_2 + \dots + x_n &= b_1 \\ x_1 + (1 + a_2)x_2 + \dots + x_n &= b_2 \\ &\vdots \\ x_1 + \dots + (1 + a_n)x_n &= b_n. \end{aligned}$$

Introducimos la incógnita auxiliar $s = x_1 + x_2 + \dots + x_n$. El sistema implica: $s + a_i x_i = b_i$ ($1 \leq i \leq n$) y pongamos

$$P = a_1 a_2 \dots a_n \quad \text{y} \quad A_i = \prod_{j \neq i} a_j.$$

Multiplicamos la i -ésima de estas relaciones por A_i y sumemos miembro a miembro:

$$s \left(P + \sum_{i=1}^n A_i \right) = \sum_{i=1}^n A_i b_i.$$

O sea

$$\Delta = P + \sum_{i=1}^n A_i = P \left(1 + \sum_{i=1}^n \frac{1}{a_i} \right).$$

Cuando $\Delta \neq 0$, s se obtiene de forma única por medio de:

$$s = \frac{\sum_{i=1}^n A_i b_i}{\Delta},$$

de donde, con la ayuda de la relación:

$$a_i x_i = b_i - s,$$

se obtiene x_i , lo cual nos demuestra que en este caso el sistema posee una solución única. Esto nos induce a pensar que Δ es el determinante del sistema, lo cual se puede ver fácilmente de forma directa.

Observemos que, en virtud de la fórmula (4), este procedimiento permite calcular la inversa de la matriz:

$$\begin{bmatrix} 1 + a_1 & 1 & 1 \\ 1 & 1 + a_2 & \\ \vdots & & \\ 1 & \dots\dots & 1 + a_n \end{bmatrix} \quad \text{cuando } \Delta \neq 0.$$

Capítulo XI

Reducción de las matrices cuadradas y aplicaciones

● En este capítulo sólo consideraremos espacios vectoriales *no nulos* sobre un cuerpo *conmutativo* K .

§ XI.1 VALORES PROPIOS, POLINOMIO CARACTERÍSTICO

DEFINICIÓN XI.1.1

Sea u un endomorfismo de un espacio vectorial E . Se dice que un elemento $x \in E$ es un **vector propio** de u si

- a) $x \neq 0$,
- b) existe $\lambda \in K$ tal que $u(x) = \lambda x$. A este elemento λ (necesariamente **único**) se le llama el **valor propio** asociado a x , y x es un **vector propio** asociado al **valor propio** λ .

El conjunto de los valores propios de un endomorfismo depende esencialmente del cuerpo de base K .

Designemos por \mathcal{I} al endomorfismo identidad. Por definición, el conjunto de los valores propios de $u \in \mathcal{L}(E)$ coincide con el conjunto de los $\lambda \in K$ para los que $u - \lambda \mathcal{I}$ no es inyectivo.

Supongamos que E es un espacio de dimensión finita n . Entonces $u - \lambda \mathcal{I}$ es *no inyectivo* si, y sólo si, es *no invertible* (esto no se verificará en dimensión infinita), o lo que es lo mismo, si

$$(1) \quad \det(u - \lambda \mathcal{I}) = 0 \quad (\text{cf. T. X.2.1}).$$

Para desarrollar (1), elegimos una base (e_1, \dots, e_n) de E . Designamos por $U = [a_{ij}]$ a la matriz de u en esta base; la matriz de $u - \lambda \mathcal{O}$ es $U - \lambda I_n$, y (1) se escribe

$$(2) \quad \det(U - \lambda I_n) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Esta relación es de la forma:

$$(-1)^n \lambda^n + \sum_{p=1}^n \alpha_p \lambda^{n-p} = 0,$$

con
$$\alpha_n = \det(U), \quad \alpha_1 = (-1)^{n-1} \left(\sum_{i=1}^n a_{ii} \right).$$

(Al elemento $\sum_i a_{ii}$ se le llama **traza** de U , y se indica por $\text{Tr}(U)$).

El primer miembro de (1) es, pues, una *función polinomio* en λ . Si K es infinito, dicha función define un polinomio formal. Pero si K es finito, no es así, por lo que vamos a definir un *polinomio formal* cuya función polinomio asociada sea $\det(u - \lambda \mathcal{O})$.

Designaremos por X cierta variable sobre K , y para toda base \mathcal{B} de E ($\mathcal{B} = (e_1, \dots, e_n)$), designaremos por $U_{\mathcal{B}}$ a la matriz de u en \mathcal{B} . En primer lugar probaremos:

El elemento $\det(U_{\mathcal{B}} - XI_n)$ de $K[X]$ depende únicamente de u (y no depende de \mathcal{B}).

En efecto, si $\mathcal{C} = (f_1, \dots, f_n)$ es otra base de E , sea P la matriz de cambio de base que pasa de \mathcal{B} a \mathcal{C} . Se tiene, $U_{\mathcal{C}} = P^{-1} U_{\mathcal{B}} P$, de donde (los cálculos se efectúan en el anillo de matrices $M_n(K(X))$ sobre el cuerpo $K(X)$):

$$U_{\mathcal{C}} - XI_n = P^{-1} U_{\mathcal{B}} P - XP^{-1} I_n P = P^{-1} (U_{\mathcal{B}} - XI_n) P,$$

luego:

$$\begin{aligned} \det(U_{\mathcal{C}} - XI_n) &= \det[P^{-1} (U_{\mathcal{B}} - XI_n) P] = \det(P^{-1}) \det(U_{\mathcal{B}} - XI_n) \det(P) \\ &= \det(P^{-1}) \det(P) \det(U_{\mathcal{B}} - XI_n) = \det(P^{-1} P) \det(U_{\mathcal{B}} - XI_n) \\ &= \det(I_n) \det(U_{\mathcal{B}} - XI_n) = \det(U_{\mathcal{B}} - XI_n). \end{aligned}$$

Podemos establecer, por lo tanto:

DEFINICIÓN XI.1.2

- $\{$ Sea u un endomorfismo de un espacio vectorial E de dimensión finita.
- $\}$ Se llama **polinomio característico de u** , y se designa por $P_u(X)$,

$\left\{ \begin{array}{l} \text{al polinomio igual, en cualquier base } \mathcal{B} \text{ de } E, \text{ a } \det(U_{\mathcal{B}} - XI_n), \text{ donde} \\ U_{\mathcal{B}} \text{ designa a la matriz de } u \text{ en } \mathcal{B}. \end{array} \right.$

Es claro que la función polinomio asociada a P_u sobre K es la aplicación $\lambda \mapsto \det(u - \lambda \mathcal{O})$; $K \rightarrow K$.

Las raíces de P_u en K son, pues, los valores propios de u . Puesto que los coeficientes del polinomio característico de u dependen únicamente de u , son invariantes de u . Por ejemplo, $\det(U)$ depende únicamente de u (cf. Cap. X), pero también $\text{Tr}(U) = \sum a_{ii}$ depende únicamente de u . A este invariante se le llama *traza del endomorfismo* u , y se indica por $\text{Tr}(u)$.

Por extensión, si U es una n -matriz cuadrada, al polinomio $\det(U - XI_n)$ se le llama *polinomio característico de* U , y a sus raíces en K se les denomina *valores propios de* U .

TEOREMA XI.1.1

$\left\| \begin{array}{l} \text{Si el cuerpo } K \text{ es algebraicamente cerrado, todo endomorfismo de un} \\ K\text{-espacio vectorial de dimensión finita } \geq 1 \text{ posee, por lo menos, un vec-} \\ \text{tor propio.} \end{array} \right.$

Demostración. El polinomio característico de todo endomorfismo tiene entonces una raíz λ en K , por lo menos, y el núcleo de $u - \lambda \mathcal{O}$ no se reduce a $\{0\}$ (cf. T. VIII.4.3).

En la práctica, XI.1.1 se aplica sobre todo a $K = \mathbb{C}$.

Triangulación

A una matriz $U = [a_{ij}]_{1 \leq i, j \leq n}$ se le llama *triangular inferior* (resp. *triangular superior*) si $a_{ij} = 0$ para $i < j$ (resp. $i > j$). El conjunto de las matrices triangulares inferiores (resp. superiores) forma una subálgebra de $M_n(K)$ (cf. § IX.1).

Sea u un endomorfismo del espacio vectorial E , de dimensión n . La matriz $U = [a_{ij}]$ asociada a u en la base (e_1, \dots, e_n) es *triangular superior* si, y sólo si, para todo $i = 1, 2, \dots, n$, el subespacio vectorial $F_i = \text{vect}(\{e_1, \dots, e_i\})$ es *estable* para u , es decir, verifica $u(F_i) \subset F_i$ (cf. § IX.2).

Para que la matriz U de u en la base (e_1, e_2, \dots, e_n) sea *triangular superior*, es necesario y suficiente que la matriz V de u en la base

$$(e_n, e_{n-1}, \dots, e_1)$$

sea *triangular inferior*, y la matriz V se deduce de U por una «simetría respecto de su centro». Las matrices U y V son semejantes, luego se deduce que *toda matriz triangular inferior es semejante a una matriz triangular superior*.

TEOREMA XI.1.2

Sea $u \in \mathcal{L}(E)$, en donde E es un espacio vectorial de dimensión finita n . Para que exista una base de E en la que la matriz de u sea triangular superior, es necesario y suficiente que $P_u(X)$ se descomponga en K en factores lineales (lo que ocurre siempre si K es algebraicamente cerrado, por lo tanto si $K = \mathbf{C}$).

Este teorema se puede enunciar en la forma equivalente:

«Para que una matriz cuadrada sea semejante a una matriz triangular inferior, es necesario y suficiente que su polinomio característico se descomponga en K en factores lineales».

En particular: toda matriz cuadrada sobre un cuerpo conmutativo algebraicamente cerrado es semejante a una matriz triangular superior (resp. inferior).

Demostración

a) Sea (e_1, \dots, e_n) una base de E en la que la matriz U de u sea triangular superior:

$$U = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & \dots & a_{22} & \dots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & a_{nn} \end{bmatrix}.$$

El polinomio $P_u(X)$ es:

$$\det(U - XI_n) = \prod_{i=1}^n (a_{ii} - X).$$

La condición es, por lo tanto, necesaria.

b) Si $P_u(X)$ se descompone en factores lineales en K , vemos, por recurrencia sobre n , que existe una base en que la matriz de u es triangular superior. La propiedad es evidente si $n = 1$. Supongámosla cierta para el entero $n - 1 \geq 1$, y demostrémosla para el entero n . Puesto que $P_u(X)$ tiene, por lo menos, una raíz λ en K , u posee un vector propio $\varepsilon_1 \neq 0$ asociado a este valor propio. Sea H un hiperplano vectorial cualquiera, suplementario de la recta $K\varepsilon_1$ engendrada por ε_1 . Si $(\varepsilon_2, \dots, \varepsilon_n)$ es una base de H , $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ es una base de E . En esta base, la matriz de u es de la forma:

$$U = \begin{bmatrix} \lambda & b_2 & \dots & b_n \\ 0 & & & \\ 0 & & V & \\ \vdots & & & \\ 0 & & & \end{bmatrix}$$

en donde V designa una matriz cuadrada de orden $n - 1$, y b_2, \dots, b_n son elementos de K .

Designemos por v el endomorfismo de H cuya matriz en $(\varepsilon_2, \dots, \varepsilon_n)$ es V . Se tiene:

$$P_u(X) = \det(U - XI_n) = (\lambda - X) \det(V - XI_{n-1}) = (\lambda - X) P_v(X);$$

lo que demuestra que $P_v(X)$ se descompone en factores lineales. Por la hipótesis de recurrencia existe una base (e_2, \dots, e_n) de H en la que la matriz de v es triangular superior.

Luego, para $i = 2, 3, \dots, n$, se tiene:

$$(3) \quad u(e_i) = b_i \varepsilon_1 + v(e_i).$$

Puesto que la matriz de v en la base (e_2, \dots, e_n) de H es triangular superior y que $u(\varepsilon_1) = \varepsilon_1$, de (3) se deduce que la matriz de u en la base $(\varepsilon_1, e_2, \dots, e_n)$ es triangular superior. c.q.d.

§ XI.2 SUBESPACIOS PROPIOS

DEFINICIÓN XI.2.1

Si u designa un endomorfismo de un espacio vectorial E y λ un valor propio de u , el **subespacio propio asociado a λ** es el conjunto E_λ ;

$$E_\lambda = \{ x \mid x \in E \text{ y } u(x) = \lambda x \} = \text{Ker}(u - \lambda \mathcal{O}).$$

E_λ está, pues, formado por el elemento 0 de E , y por los vectores propios de u asociados al valor propio λ . Este es un subespacio de E y puesto que, por definición, un vector propio es no nulo, y se tiene siempre $\dim(E_\lambda) \geq 1$.

TEOREMA XI.2.1

Sean E un espacio vectorial de dimensión n , y u un endomorfismo de E ; designemos por $\lambda_1, \dots, \lambda_p$ a los valores propios **distintos** de u , y por E_1, \dots, E_p a los subespacios propios correspondientes. Entonces la suma $E_1 + E_2 + \dots + E_p$ es directa (cf. Cap. VIII. § 1, def. VIII.1.3).

Demostración. Por reducción al absurdo. Si la suma $E_1 + E_2 + \dots + E_p$ no fuese directa, existirían relaciones de la forma:

$$(1) \quad \sum_{i=1}^p \mu_i x_i = 0,$$

con $x_i \in E_i - \{0\}$, y los μ_i no todos nulos.

Consideremos una relación de la forma (1), en que el número r de escalares μ_i no nulos sea *mínimo* (nota: necesariamente $r \geq 2$). Podemos suponer, entonces, que dicha relación es:

$$(2) \quad v_1 e_1 + v_2 e_2 + \dots + v_r e_r = 0, \quad e_i \in E_i \setminus \{0\}; \quad v_1 \neq 0, v_2 \neq 0, \dots, v_r \neq 0.$$

Sea v el primer miembro de (2). Formamos $u(v) - \lambda_1 v$, y se obtiene:

$$u(v) - \lambda_1 v = (\lambda_2 - \lambda_1) v_2 e_2 + \dots + (\lambda_r - \lambda_1) v_r e_r = 0.$$

Los $r - 1$ coeficientes $(\lambda_k - \lambda_1) v_k$ ($2 \leq k \leq r$) son no nulos, lo que contradice el carácter minimal de (2). c.q.d.

DEFINICIÓN XI.2.2

§ Un endomorfismo u del espacio vectorial E es **diagonalizable** si E es suma directa de los subespacios propios de u .

Si E es de dimensión finita, equivale a decir que E admite una base formada por vectores propios de u . En dicha base, la matriz de u es de la forma:

$$(3) \quad \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

en donde $\lambda_1, \dots, \lambda_n$ son los valores propios de u (cada uno de ellos repetido tantas veces como indique su multiplicidad). La matriz (3) es *diagonal*, de ahí la terminología empleada.

Sea U la matriz de u en una base cualquiera de E , si Q designa la matriz del cambio de base que transforma esta base en la base de los vectores propios, vemos que la matriz (3) es igual a $Q^{-1} U Q$. Por lo tanto, u es diagonalizable si, y sólo si, su matriz en una base cualquiera de E es semejante a una matriz diagonal.

Por abuso de lenguaje, se dice también que una matriz cuadrada es diagonalizable si es semejante a una matriz diagonal.

El teorema XI.1.1 nos da inmediatamente el

TEOREMA XI.2.2

|| Sea u un endomorfismo de un espacio vectorial E de dimensión n . Si u admite n valores propios distintos, u es diagonalizable.

Demostración. Sean E_1, \dots, E_n los n subespacios propios. Se tiene, para todo i , $\dim(E_i) \geq 1$. Y en virtud del teorema XI.2.1, la suma $E_1 + \dots + E_n$ es directa, de donde se sigue que $\sum_{i=1}^n \dim(E_i) \leq n$, y finalmente $\dim(E_i) = 1$ para todo i , lo que implica el resultado enunciado. c.q.d.

Nota. El teorema XI.2.2 da una condición *suficiente* para que un endomorfismo sea diagonalizable. Más adelante veremos que esta condición *no es necesaria*.

En los §§ que siguen se realizará un estudio más profundo de los subespacios propios. De momento, nos limitamos a describir un primer método para hallar la dimensión de los subespacios propios de un endomorfismo.

Designemos por λ un valor propio del endomorfismo u , y en una base fija, sea $U = [a_{ij}]$ la matriz de u . Los vectores propios de u asociados a λ son aquellos cuyas coordenadas (x_1, x_2, \dots, x_n) verifican el sistema lineal

$$(4) \quad (U - \lambda I_n) \mathcal{X} = 0, \quad \text{con } \mathcal{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

En virtud de los resultados obtenidos acerca de los sistemas lineales (§ X.5) podemos enunciar el

TEOREMA XI.2.3

Sean E un espacio vectorial de dimensión n , u un endomorfismo de E , y U la matriz de u en una base cualquiera. Entonces la **dimensión** del subespacio propio E_λ asociado a un valor propio λ de u , es igual a $n - r$, en donde r designa el **rango** de la matriz $U - \lambda I_n$.

Con las notaciones precedentes, sean u un endomorfismo del espacio vectorial E , de valores propios distintos $\lambda_1, \dots, \lambda_p$, E_1, \dots, E_p los subespacios propios asociados, y $\beta_i = \dim(E_i)$. El polinomio característico de u es de la forma

$$P_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i} Q(X), \quad \text{con } \sum_{i=1}^p \alpha_i \leq n = \dim(E),$$

en donde $Q(X)$ designa un polinomio sin raíces en K .

Al entero α_i se le llama *multiplicidad* del valor propio λ_i . Vamos a demostrar que se verifica siempre $\beta_i \leq \alpha_i$. (El teorema XI.2.1 implica ya que $\sum \beta_i \leq n$.)

TEOREMA XI.2.4

Sean E un espacio vectorial de dimensión n , u un endomorfismo de E , λ un valor propio de u de multiplicidad α , E_λ el subespacio propio asociado a λ . Se tiene:

$$\dim(E_\lambda) \leq \alpha.$$

Demostración. Por reducción al absurdo. Supongamos que es $\dim(E_\lambda) \geq \alpha + 1$. Entonces existe una base (e_1, \dots, e_n) de E tal que $e_1 \in E_\lambda, \dots, e_{\alpha+1} \in E_\lambda$. En esta base, la matriz de u tiene la forma:

$$\begin{array}{c} \alpha + 1 \text{ filas} \left\{ \begin{array}{c} \overbrace{\alpha + 1 \text{ columnas}} \\ \left[\begin{array}{cccc|ccc} \lambda & 0 & \dots & 0 & \times & \dots & \times \\ & \ddots & & \vdots & & & \vdots \\ 0 & & \ddots & 0 & & & \vdots \\ & & & \ddots & & & \vdots \\ & & & & 0 & & \vdots \\ 0 & \dots & 0 & \lambda & \times & & \vdots \\ & & & & & \times & \vdots \\ 0 & \dots & \dots & 0 & \times & & \vdots \\ & & & & & \ddots & \vdots \\ & & & & & & \times \end{array} \right] \end{array} \right. \end{array}$$

(en donde los signos \times designan escalares cualesquiera).

Si calculamos $P_u(X)$ en esta base, vemos que $(\lambda - X)^{\alpha+1}$ es un factor de $P_u(X)$, lo cual es absurdo. c.q.d.

COROLARIO

Sean E un espacio vectorial de dimensión n , y $u \in \mathcal{L}_K(E)$. Para que u sea diagonalizable, es necesario y suficiente que se verifiquen las dos condiciones siguientes:

- a) El polinomio característico $P_u(X)$ se descompone en factores lineales en $K[X]$.
- b) La dimensión de cualquiera de los subespacios propios de u es igual a la multiplicidad del valor propio asociado.

Demostración. Si a) y b) se verifican, el que u sea diagonalizable es consecuencia elemental de XI.2.1.

Recíprocamente, supongamos que u es diagonalizable, y sean $(\lambda_1, \dots, \lambda_p)$ sus valores propios distintos, α_i la multiplicidad de λ_i , y E_i el subespacio propio asociado a λ_i ($1 \leq i \leq p$).

Se tiene, según XI.2.4:

$$\forall i \dim(E_i) \leq \alpha_i, \text{ de donde } \sum_{i=1}^p \dim(E_i) \leq \sum_{i=1}^p \alpha_i \leq n.$$

Pero, al ser E suma directa de los E_i , se tiene también $\sum_{i=1}^p \dim(E_i) = n$. De donde $\sum_{i=1}^p \alpha_i = n$, lo que prueba que $P_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$. Este polinomio se descompone, pues, en factores lineales en K . Además, necesariamente es $\forall i \dim(E_i) = \alpha_i$, en virtud de las desigualdades $\dim(E_i) \leq \alpha_i$ y de $\sum_i \dim(E_i) = \sum_i \alpha_i$. c.q.d.

Ejemplos

1) Valores propios y vectores propios de A en $M_4(\mathbf{C})$.

$$A = \begin{bmatrix} 1-it & 0 & 2it & 0 \\ 0 & 1-it & 2it & 0 \\ 0 & 0 & 1+it & 0 \\ 0 & 0 & 0 & 1+it \end{bmatrix} \quad t, \text{ real no nulo.}$$

Se considera A como matriz de un endomorfismo u de \mathbf{C}^4 en la base canónica. Los dos valores propios dobles, $\lambda = 1 + it$ y $\bar{\lambda} = 1 - it$, son evidentes. La matriz $A - \lambda I_4$ es

$$\begin{bmatrix} -2it & 0 & 2it & 0 \\ 0 & -2it & 2it & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{y su rango es 2.}$$

El subespacio propio E_λ es, pues, de dimensión 2.

Una base de E_λ está formada por los vectores v_1, v_2 de coordenadas $(1, 1, 1, 1)$ y $(1, 1, 1, 0)$. Análogamente:

$$A - \lambda I_4 = \begin{bmatrix} 0 & 0 & 2it & 0 \\ 0 & 0 & 2it & 0 \\ 0 & 0 & 2it & 0 \\ 0 & 0 & 0 & 2it \end{bmatrix} \quad \text{es de rango 2.}$$

El subespacio propio $E_{\bar{\lambda}}$ es, pues, de dimensión 2.

Una base de $E_{\bar{\lambda}}$ está formada por los vectores v_3, v_4 de coordenadas $(1, 1, 0, 0)$ y $(1, 0, 0, 0)$. De todo ello se deduce que A es diagonalizable; en la base (v_i) el endomorfismo u se halla representado por la matriz

$$D = \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \bar{\lambda} & 0 \\ 0 & 0 & 0 & \bar{\lambda} \end{bmatrix}.$$

La matriz de cambio de base que pasa de la base canónica a la base (v_i) es:

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Un cálculo fácil demuestra que:

$$P^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix}.$$

(En efecto, el endomorfismo asociado a P está dado por las fórmulas

$$y_1 = x_1 + x_2 + x_3 + x_4, \quad y_2 = x_1 + x_2 + x_3, \quad y_3 = x_1 + x_2, \quad y_4 = x_1,$$

de donde se deducen las fórmulas inversas:

$$x_1 = y_4, \quad x_2 = y_3 - y_4, \quad x_3 = y_2 - y_3, \quad x_4 = y_1 - y_2.)$$

Se tiene, pues: $D = P^{-1}AP$.

A modo de aplicación, vamos a calcular la potencia n -ésima de la matriz A . A^n es la matriz de u^n en la base canónica. En la base (v_i) , la matriz de u^n es D^n ; de donde:

$$(5) \quad A^n = P^{-1} \cdot D^n \cdot P;$$

puesto que el cálculo de D^n es inmediato, esta fórmula nos da A^n . De hecho, resulta inútil realizar el cálculo. Observemos que A se escribe:

$$A = \begin{bmatrix} \bar{\lambda} & 0 & \lambda - \bar{\lambda} & 0 \\ 0 & \bar{\lambda} & \lambda - \bar{\lambda} & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix};$$

luego, D^n tiene la misma forma que D , con la diferencia de que λ se ha substituido por λ^n . Puesto que A^n es la matriz semejante a D^n , en el cambio de base definido por los (v_i) , vemos que A^n se obtiene simplemente cambiando λ por λ^n (por lo tanto, $\bar{\lambda}$ por $\bar{\lambda}^n$)

$$A^n = \begin{bmatrix} (1-it)^n & 0 & (1+it)^n - (1-it)^n & 0 \\ 0 & (1-it)^n & (1+it)^n - (1-it)^n & 0 \\ 0 & 0 & (1+it)^n & 0 \\ 0 & 0 & 0 & (1+it)^n \end{bmatrix}.$$

2) Incluso en el caso de que una matriz sea diagonalizable, la fórmula (5) no es siempre el método más rápido para calcular su potencia n -ésima.

Por ejemplo, consideremos la matriz:

$$J = [a_i b_j]_{1 \leq i, j \leq n}, \quad (a_i, b_j \in R)$$

en donde a_1, \dots, a_n son no todos nulos, y b_1, \dots, b_n tampoco son todos nulos. Un cálculo elemental prueba

$$J^2 = \lambda J, \text{ en donde } \lambda = \sum_{i=1}^n a_i b_i, \text{ y por recurrencia } J^m = \lambda^{m-1} J.$$

Vamos a estudiar ahora los subespacios propios de J .

El endomorfismo u asociado a J en \mathbf{R}^n está definido por las fórmulas:

$$(6) \quad y_i = \sum_{j=1}^n a_i b_j x_j = a_i \sum_{j=1}^n b_j x_j, \text{ con } u(x) = y, \quad x = (x_i), \quad y = (y_i).$$

Puesto que los (a_i) no son todos nulos, el subespacio propio de u relativo al valor propio $\lambda = 0$ (que es el núcleo de u) es el hiperplano vectorial H de ecuación

$$\sum_{j=1}^n b_j x_j = 0.$$

(6) muestra además que la imagen de u es el subespacio de dimensión 1 engendrado por el vector $a = (a_i)$. a es vector propio de u , para el valor propio $\lambda = \sum_{i=1}^n a_i b_i$.

a) Si $\lambda \neq 0$, u es diagonalizable. Una base de vectores propios estará formada por una base cualquiera de H , y por el vector a . El polinomio característico es $(-1)^n X^{n-1} (X - \lambda)$ (en efecto, λ es valor propio no nulo, y dado que $\dim(H) = n-1$, la multiplicidad del valor propio 0 es $\geq n-1$).

b) Si $\lambda = 0$, el único valor propio es 0, el polinomio característico es $(-1)^n X^n$. La matriz no es diagonalizable, si no, representaría necesariamente el endomorfismo nulo. (Su forma diagonal sería, en efecto, la matriz nula.)

● *Nota.* En general, toda matriz no nula cuyo único valor propio sea 0 no es diagonalizable, en virtud del razonamiento anterior.

3) He aquí otro ejemplo en que es posible calcular A^p sin diagonalizar la matriz A

$$A = \begin{bmatrix} a & b & \dots & b \\ b & a & \dots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & \dots & b & a \end{bmatrix} \quad A \in M_n(K).$$

Si designamos por J_n a la matriz que tiene todos sus elementos iguales a 1 y por I_n a la matriz identidad, resulta:

$$A = (a - b) I_n + b J_n.$$

Puesto que $I_n J_n = J_n I_n = J_n$, se puede calcular A^p con la fórmula del binomio:

$$A^p = [(a - b) I_n + b J_n]^p = \sum_{q=0}^p \binom{p}{q} (a - b)^{p-q} b^q J_n^q.$$

Si tenemos en cuenta $J_n^2 = nJ$, $J_n^q = n^{q-1} J_n$, ($q \geq 1$) resulta:

$$A^p = (a - b)^p I_n + \sum_{q=1}^p \binom{p}{q} (a - b)^{p-q} b^q n^{q-1} J_n.$$

§ XI.3 POLINOMIOS DE ENDOMORFISMOS. TEOREMA DE HAMILTON-CAYLEY

Recordemos que un subespacio F del espacio vectorial E es estable para el endomorfismo u de E si se verifica $u(F) \subset F$.

Para todo polinomio $M \in K[X]$, con $M = a_m X^m + a_{m-1} X^{m-1} + \dots + a_0$, consideremos el endomorfismo $M(u)$:

$$M(u) = a_m u^m + a_{m-1} u^{m-1} + \dots + a_1 u + a_0 \mathcal{O},$$

en donde \mathcal{O} designa el endomorfismo identidad (por convenio, $u^0 = \mathcal{O}$).

La aplicación $\chi: M \mapsto M(u)$ es un homomorfismo de la K -álgebra $K[X]$ en $\mathcal{L}_K(E)$; con otras palabras, se tiene

$$(MN)(u) = M(u) \circ N(u) \quad \text{y} \quad (M + N)(u) = M(u) + N(u),$$

para todos los $M, N \in K[X]$.

DEFINICIÓN XI.3.1

$\{$ Con las notaciones anteriores, a la imagen de $K[X]$ por la aplicación χ
 $\{$ se le llama **álgebra de polinomios del endomorfismo u** , y se de-
 $\}$ signa por $K[u]$. Al homomorfismo χ se le llama **canónico**.

$K[X]$ es un álgebra conmutativa, y $K[u]$ es la imagen de esta K -álgebra por medio del homomorfismo χ . Vemos que $K[u]$ es una sub- K -álgebra *conmutativa* de $\mathcal{L}_K(E)$. Se comprueba, además, sin ninguna dificultad, la relación

$$M(u) \circ N(u) = N(u) \circ M(u),$$

si partimos del hecho (que se demuestra por recurrencia) de que todas las potencias de u conmutan entre sí.

El núcleo del homomorfismo

$$\chi: K[X] \rightarrow \mathcal{L}_K(E)$$

$$M \mapsto M(u)$$

es un ideal α_u de $K[X]$. Observemos que, si E es de dimensión finita n , el ideal α_u es *no nulo*, pues $\mathcal{L}_K(E)$ tiene entonces dimensión n^2 . Luego los endomorfismos $\mathcal{O}, u, u^2, \dots, u^{n^2}$ (en número $n^2 + 1$) están ligados, lo que significa que α_u contiene un polinomio no nulo de grado $\leq n^2$.

En lo sucesivo, *supondremos que E es de dimensión finita*.

Sabemos (T. IV.2.2) que todo ideal de $K[X]$ es un ideal *principal* (cf. Cap. III). Luego, existe un polinomio q_u , único salvo para un factor de proporcionalidad, tal que $\alpha_u = (q_u)$, y $q_u \neq 0$ ya que $\alpha_u \neq \{0\}$.

DEFINICIÓN XI.3.2

Sean E un espacio vectorial de dimensión finita, u un endomorfismo de E , y $\chi : K[X] \rightarrow \mathcal{L}_K(E)$ el homomorfismo canónico. Se llama **polinomio minimal** de u al polinomio normalizado q_u tal que el núcleo de χ es el ideal (q_u) de $K[X]$.

En otras palabras: q_u es el polinomio normalizado de grado mínimo tal que $q_u(u) = 0$. Este polinomio está unívocamente determinado. Hemos visto ya que $q_u \neq 0$ y que $\text{gr}(q_u) \leq n^2$. Vamos a precisar estos resultados.

TEOREMA XI.3.1 (Hamilton-Cayley)

Sean E un espacio vectorial de dimensión finita, u un endomorfismo de E , P_u su polinomio característico y q_u su polinomio minimal. Entonces q_u es un divisor de P_u . En otras palabras, se tiene la relación:

$$(1) \quad P_u(u) = 0.$$

Demostración. Sea $A = [a_{ij}]_{1 \leq i, j \leq n}$ la matriz de u en una base cualquiera, fijada de antemano, (e_1, \dots, e_n) . Designamos por $c_{ij}(X)$ a los adjuntos de la matriz

$$B(X) = A - XI_n$$

(con coeficientes en $K[X]$) definida por

$$B(X) = [b_{ij}(X)], \quad \text{con} \quad b_{ij}(X) = a_{ij} - \delta_{ij} X \quad (i, j = 1, 2, \dots, n),$$

en donde δ_{ij} es el símbolo de Kronecker.

Los $c_{ij}(X)$ son polinomios de grado $n - 1$ en X , y se tiene

$$\det(B(X)) = P_u(X).$$

Según las observaciones del final del § X.2, estos polinomios verifican formalmente las n^2 relaciones

$$\sum_{k=1}^n b_{ik}(X) c_{jk}(X) = \delta_{ij} P_u(X)$$

que expresan la igualdad de las matrices:

$$B(X) {}^t C(X) = \det(B(X)) I_n.$$

El cuerpo K es conmutativo, luego podemos (cambiando i y j) escribir estas relaciones en la forma

$$(2) \quad \sum_{k=1}^n c_{ik}(X) b_{jk}(X) = \delta_{ji} P_u(X).$$

Las relaciones (2) son identidades formales de polinomios, luego podemos reemplazar la variable X por el endomorfismo u (cf. principio de este §). Se tiene entonces:

$$b_{ij}(u) = a_{ij} u^0 - \delta_{ij} u = a_{ij} \mathcal{O} - \delta_{ij} u;$$

y se obtiene:

$$(3) \quad \sum_{k=1}^n c_{ik}(u) \cdot b_{jk}(u) = \begin{cases} 0 & \text{si } i \neq j \\ P_u(u) & \text{si } i = j. \end{cases}$$

El primer miembro de (3) designa una suma de productos de endomorfismos de E .

A fin de aligerar la escritura, designamos por $M(u) \cdot e$ a la imagen del vector e por el endomorfismo $M(u)$ (estas notaciones son las de la teoría de operadores y se utilizarán en el capítulo XII). Con estas notaciones, para todo vector $e \in E$, (3) implica:

$$(4) \quad \sum_{k=1}^n c_{ik}(u) \cdot [b_{jk}(u) \cdot e] = \delta_{ij} P_u(u) \cdot e.$$

Luego, por la definición de la matriz de u , se tiene:

$$u \cdot e_k = \sum_{j=1}^n a_{jk} e_j,$$

por lo tanto

$$(5) \quad \sum_{j=1}^n b_{jk}(u) \cdot e_j = \sum_{j=1}^n (a_{jk} \mathcal{O} - \delta_{jk} u) \cdot e_j = \left(\sum_{j=1}^n a_{jk} e_j \right) - u \cdot e_k = 0.$$

En (4) fijamos el índice i y sumamos miembro a miembro las relaciones obtenidas para $j = 1, 2, \dots, n$ reemplazando cada vez e por e_j en la ecuación de índice j . Utilizando (5) se obtiene:

$$\sum_{j=1}^n \delta_{ij} P_u(u) \cdot e_j = 0, \quad \text{es decir} \quad P_u(u) \cdot e_i = 0.$$

Esta última relación es verdadera para $i = 1, 2, \dots, n$, por lo tanto vemos que $P_u(u)$ es el operador nulo. c.q.d.

Otra demostración. Con las mismas notaciones, podemos escribir

$${}^tC(X) = \sum_{k=0}^{n-1} X^k C_k,$$

en donde C_0, C_1, \dots, C_{n-1} son elementos de $M_n(K)$.

Las relaciones (2) se expresan entonces por la igualdad formal (entre matrices con coeficientes en $K[X]$):

$$(5') \quad (A - XI_n) \cdot \left(\sum_{k=0}^{n-1} X^k C_k \right) = P_u(X) I_n.$$

Poniendo $P_u(X) = \sum_{k=0}^n p_k X^k$, deducimos, identificando los términos del mismo grado en X en el anillo $M_n(K[X])$

$$AC_0 = p_0 I_n, \quad AC_1 - C_0 = p_1 I_n \dots AC_k - C_{k-1} = p_k I_n \dots - C_{n-1} = p_n I_n.$$

(Estas relaciones expresan simplemente, en forma condensada, las identidades (2).) Se tiene, pues:

$$\begin{aligned} p_u(A) &= \sum_{k=0}^n p_k A^k = \sum_{k=0}^n p_k I_n \cdot A^k \\ &= AC_0 + (AC_1 - C_0) A + \dots + (AC_k - C_{k-1}) A^k + (AC_{k+1} - C_k) A^{k+1} \\ &\quad + \dots + (AC_{n-1} - C_{n-2}) A^{n-1} - C_{n-1} A^n = 0 \end{aligned}$$

(los términos escritos se destruyen dos a dos).

Por lo tanto, se tiene $p_u(A) = 0$, que equivale a $p_u(u) = 0$. c.q.d.

Notas

1) Podría tentarnos reemplazar directamente X por A en (5'). La justificación de este procedimiento es, sin embargo, delicada (haría falta utilizar el hecho de que A conmuta con las matrices C_i).

2) Si K es algebraicamente cerrado, es posible dar una demostración elegante de XI.3.1, fundamentada en una aplicación cuidadosa del teorema IV.7.5 (cf. ejercicio XI.34). En este caso, es posible elegir una base (e_i) de E en la que la matriz A de u es triangular, y P_u de la forma

$$P_u(X) = \prod_{i=1}^n (\lambda_i - X).$$

Se obtiene entonces otra demostración de XI.3.1 realizando los productos de las matrices $\lambda_i I_n - A$ y comprobando que

$$P_u(A) = \prod_{i=1}^n (\lambda_i I_n - A) = 0.$$

Extensión. Nuestras primeras demostraciones utilizan únicamente la estructura de *anillo conmutativo unífero* de K . Podemos, pues, enunciar, en forma matricial, el teorema siguiente, que es más general:

XI.3.2 Sea $A = [a_{ij}]_{1 \leq i \leq n, 1 \leq j \leq n}$ una matriz cuadrada con elementos en un **anillo conmutativo unífero** K . Designemos por $P_A(X)$ al polinomio $\det(A - XI_n)$, con coeficientes en K (polinomio característico de A). Entonces se tiene:

$$P_A(A) = 0.$$

Para terminar este §, damos algunos resultados utilizados en el estudio de los subespacios característicos de un endomorfismo, que constituirá el objeto del § 4:

TEOREMA XI.3.3

Sean E un K -espacio vectorial de dimensión finita n , y u un endomorfismo de E . Si S_1, S_2, \dots, S_p designan **polinomios primos entre sí dos a dos**, con coeficientes en K , y $\sigma_1, \sigma_2, \dots, \sigma_p$ los operadores $S_1(u), S_2(u), \dots, S_p(u)$, el núcleo N de $\sigma = \sigma_1 \sigma_2 \dots \sigma_p$ es la **suma directa** de los núcleos N_1, N_2, \dots, N_p de $\sigma_1, \sigma_2, \dots, \sigma_p$.

Demostración. El teorema es evidente si $p = 1$. Para $p \geq 2$, razonamos por recurrencia sobre p .

a) Caso en que $p = 2$.

Según el teorema de Bezout, existen polinomios $U_1, U_2 \in K[X]$ tales que $U_1 S_1 + U_2 S_2 = 1$, de donde, designando por u_i al operador $U_i(u)$ y por \mathcal{O} al operador identidad

$$(6) \quad u_1 \cdot \sigma_1 + u_2 \cdot \sigma_2 = \mathcal{O}.$$

Luego, para todo $x \in E$, se tiene:

$$(7) \quad (u_1 \cdot \sigma_1) \cdot x + (u_2 \cdot \sigma_2) \cdot x = x.$$

Si $x \in N$, se tiene (utilizando la conmutatividad de $K[u]$):

$$[\sigma_2 \cdot (u_1 \cdot \sigma_1)] \cdot x = (u_1 \cdot \sigma_1 \cdot \sigma_2) \cdot x = u_1(\sigma \cdot x) = 0,$$

de donde $(u_1 \cdot \sigma_1) \cdot x \in N_2$, y análogamente $(u_2 \cdot \sigma_2) \cdot x \in N_1$. Luego $N \subset N_1 + N_2$. Por otro lado $N_1 \subset N$ y $N_2 \subset N$, como la relación $\sigma_1 \cdot x = 0$ implica

$$\sigma_2(\sigma_1 \cdot x) = \sigma \cdot x = 0,$$

y también $\sigma_2 \cdot x = 0$ implica $\sigma \cdot x = 0$, se deduce que $N = N_1 + N_2$.

Finalmente (7) muestra inmediatamente que $N_1 \cap N_2 = \{0\}$, de donde se obtiene el teorema cuando $p = 2$.

b) Supongamos el teorema verdadero para el entero $p - 1 \geq 2$, y demostremos que también es verdadero para el entero p .

Para ello hacemos $S = S_1 S_2 \dots S_p$, $R = S_1 S_2 \dots S_{p-1}$, y designamos por M al núcleo del endomorfismo $R(u)$. Los polinomios R y S_p son primos entre sí, luego (según la parte a) de la demostración) N es la suma directa de M y de N_p . Pero según la hipótesis de recurrencia, M es la suma directa de N_1, \dots, N_{p-1} , y según una propiedad evidente de asociatividad de la suma directa, se deduce que N es la suma directa de N_1, N_2, \dots, N_p . c.q.d.

COROLARIO

Sean S_1, S_2, \dots, S_p p polinomios primos entre sí, dos a dos. Con las notaciones de XI.3.3, supongamos que $\sigma_1 \sigma_2 \dots \sigma_p = 0$. Entonces E es la suma directa de los núcleos N_i de los operadores σ_i ($1 \leq i \leq p$).

De lo que precede se deduce inmediatamente el teorema fundamental siguiente:

TEOREMA XI.3.4

Sea u un endomorfismo de un K -espacio vectorial E de dimensión finita; y sea $Q(X) = \prod_{i=1}^p (\lambda_i - X)^{v_i}$ un polinomio descompuesto en factores lineales sobre el cuerpo K , tal que $Q(u) = 0$ (los λ_i se suponen distintos dos a dos). Entonces E es suma directa de los núcleos de los endomorfismos $(\lambda_i \mathcal{O} - u)^{v_i}$.

§ XI.4 SUBESPACIOS CARACTERÍSTICOS

● En lo que sigue, el cuerpo de base K se supondrá algebraicamente cerrado.

De lo que resulta que el polinomio característico del endomorfismo u se escribe

$$P_u(X) = \prod_{i=1}^n (\lambda_i - X)^{\alpha_i},$$

en donde $\lambda_1, \dots, \lambda_p$ designan los valores propios *distintos* de u , y α_i el orden de la multiplicidad del valor propio λ_i .

Del teorema XI.3.1 (Hamilton-Cayley) resulta inmediatamente que el polinomio minimal de u se escribe entonces

$$q_u(X) = \prod_{i=1}^p (X - \lambda_i)^{\beta_i} \quad (0 \leq \beta_i \leq \alpha_i).$$

Vamos a precisar esta situación.

DEFINICIÓN XI.4.1

Sea E un espacio vectorial de dimensión finita, y $u \in \mathcal{L}(E)$ un endomorfismo de E , de polinomio característico $P_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$ (los λ_i son los valores propios, distintos, de u).
Al núcleo del endomorfismo $(\lambda_i \mathcal{O} - u^{\alpha_i})$ se le llama **subespacio característico** de u , asociado al valor propio λ_i .

Con esta definición, los teoremas XI.3.1 y XI.3.4 implican:

XI.4.1 Cualquiera que sea el endomorfismo u del espacio vectorial E (de dimensión finita, sobre un cuerpo algebraicamente cerrado) el espacio E es suma directa de los subespacios característicos de u .

Designemos por N_i el núcleo de $(\lambda_i \mathcal{O} - u)^{\alpha_i}$. El subespacio propio E_i asociado al valor propio λ_i es el núcleo de $\lambda_i \mathcal{O} - u$. Se tiene pues $E_i \subset N_i$.

Además, $(\dim(E_i) \geq 1) \Rightarrow (\dim(N_i) \geq 1)$. Esta propiedad tiene una consecuencia importante:

XI.4.2 Sea $q_u(X) = \prod_{i=1}^p (X - \lambda_i)^{\beta_i}$ el polinomio minimal de u , en donde $\lambda_1, \dots, \lambda_p$ designan los valores propios distintos de u ; entonces para $1 \leq i \leq p$, se tiene $\beta_i \geq 1$. En otras palabras, todo valor propio de u es raíz de su polinomio minimal. Además, el subespacio característico relativo al valor propio λ_i es también el núcleo de $(\lambda_i \mathcal{O} - u)^{\beta_i}$.

Demostración. Designemos por M_i el núcleo de $(\lambda_i \mathcal{O} - u)^{\beta_i}$. Según XI.3.3, como $q_u(u) = 0$, E es la suma directa de los M_i que no son nulos, es decir, de los M_i que corresponden a los valores de i para los cuales $\beta_i > 0$. Si $P_u(X) = \prod_{i=1}^p (\lambda_i -$

$-X)_{ai}$ es el polinomio característico de u , y si $N_i = \text{Ker}(\lambda_i \mathcal{D} - u)^{a_i}$ designa el subespacio característico asociado a λ_i , se tiene evidentemente $M_i \subset N_i$, ya que $\beta_i \leq a_i$ (cf. XI.3.1). Dado que E es suma directa de los N_i , se deduce que $M_i = N_i$ para todo i , luego $M_i \neq \{0\}$ y $\beta_i \geq 1$ para todo $i = 1, 2, \dots, p$. c.q.d.

Nota 1. En general, si γ_i es un entero $\geq \beta_i$, el subespacio característico N_i es el núcleo de $(\lambda_i \mathcal{D} - u)^{\gamma_i}$.

Nota 2. He aquí una demostración directa de XI.4.2. Designamos por $P_u(X)$ y por $q_u(X)$ al polinomio característico y al polinomio minimal de u .

Ponemos $q_u(X) = X^k + a_1 X^{k-1} + \dots + a_k$. Si λ es valor propio de u , λ^k es valor propio de u^k para todo $k \in \mathbf{N}$. Se tiene:

$$0 = q_u(u) = u^k + a_1 u^{k-1} + \dots + a_k \mathcal{D};$$

y si x designa a un vector propio de u asociado a λ , se deduce:

$$0 = q_u(u) \cdot x = (\lambda^k + a_1 \lambda^{k-1} + \dots + a_k) x = q_u(\lambda) x.$$

Puesto que x es no nulo, resulta $q_u(\lambda) = 0$ (puesto que $q_u(\lambda)$ es un escalar).

Propiedades de los subespacios característicos

Conservamos las notaciones anteriores. Ante todo vemos que *cada uno de los subespacios característicos N_i es estable para u* . En efecto $(\lambda_i \mathcal{D} - u)^{\beta_i} \cdot u(x) = [u \cdot (\lambda_i \mathcal{D} - u)^{\beta_i}] \cdot x$, luego

$$((\lambda_i \mathcal{D} - u)^{\beta_i} \cdot x = 0) \Rightarrow ((\lambda_i \mathcal{D} - u)^{\beta_i} \cdot u(x) = 0) \quad \text{y} \quad (x \in N_i) \Rightarrow (u(x) \in N_i).$$

La restricción de u a N_i , a saber u_i , es entonces un endomorfismo de N_i . Por definición, $(\lambda_i \mathcal{D} - u_i)^{\beta_i} = 0$. Luego el polinomio minimal q_i de u_i divide a $(\lambda_i - X)^{\beta_i}$, de donde $q_i = (X - \lambda_i)^{\gamma_i}$, con $\gamma_i \leq \beta_i$. Si fuera $\gamma_i < \beta_i$, el polinomio $r = q_i \prod_{j \neq i} (X - \lambda_j)^{\beta_j}$, de grado estrictamente inferior al de q_u verificaría $r(u) = 0$, y q_u no sería el polinomio minimal. En resumen:

XI.4.3 Si u_i designa la restricción del endomorfismo u al subespacio característico E_i , el polinomio minimal de u_i es $(X - \lambda_i)^{\beta_i}$ (con las mismas notaciones que en XI.4.2) ⁽¹⁾.

⁽¹⁾ En particular, $\beta_i = 1$ sí, y sólo, si la restricción de u a N_i es una homotecia (cf. la nota que sigue a la proposición XI.4.2), en otras palabras, si N_i es igual al subespacio propio relativo a λ_i .

TEOREMA XI.4.4

Sean E un espacio vectorial de dimensión n , $u \in \mathcal{L}(E)$ un endomorfismo de E , $P_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$ su polinomio característico (en donde $\lambda_1, \dots, \lambda_p$ son los valores propios distintos de u). Si N_i ($i = 1, 2, \dots, p$) designa el subespacio característico asociado a λ_i , se tiene $\dim(N_i) = \alpha_i$.

Demostración. Elijamos una base adaptada a la descomposición de E en suma directa de los N_i , es decir de la forma:

$$(e_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq v_i}}, \text{ en donde } v_i = \dim(N_i),$$

y en donde, para todo i , $(e_{ij})_{1 \leq j \leq v_i}$ sea una base de N_i . En esta base, puesto que cada N_i es estable para u la matriz U de u se escribe,

$$(4) \quad U = \begin{bmatrix} \boxed{U_1} & 0 & \dots & 0 \\ 0 & \boxed{U_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ \vdots & & & \boxed{U_i} & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \dots & 0 & & & \boxed{U_p} \end{bmatrix};$$

U_i designa una matriz cuadrada de orden v_i , que es la de la restricción u_i de u a N_i . La diagonal principal de U_i está incluida en la de U . Antes hemos visto que el único valor propio de U_i es λ_i , ya que el polinomio característico de U_i es:

$$P_i = (\lambda_i - X)^{v_i}.$$

Pero (corolario del teorema X.2.4):

$$\det(U - XI_n) = \prod_{i=1}^p \det(U - XI_{v_i}),$$

luego el polinomio característico P de U es

$$\prod_{i=1}^p (\lambda_i - X)^{v_i}.$$

Dado que los λ_i son distintos, y que $P = P_u = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$, tenemos para $1 \leq i \leq p$

$$v_i = \alpha_i. \text{ c.q.d.}$$

La reducción de la matriz del endomorfismo u a la forma (4), por un cambio de base adecuado, es una *reducción de u según sus subespacios característicos*. En la mayor parte de las aplicaciones prácticas, esta reducción es suficiente.

Aplicación a las ecuaciones diferenciales (ver tomo 4 de la presente obra)

Aplicación a las sucesiones recurrentes

Sean $\alpha_0, \alpha_1, \dots, \alpha_{k-1}$ números complejos fijos ($\alpha_0 \neq 0$). Buscamos las sucesiones $(u_n)_{n \geq 0}$ de números complejos que verifiquen, para todo $n \in \mathbf{N}$, la relación

$$(5) \quad u_{n+k} = \alpha_{k-1} u_{n+k-1} + \dots + \alpha_0 u_n.$$

El conjunto \mathcal{E} de las sucesiones que verifican (5) es un subespacio del espacio $\mathbf{C}^{\mathbf{N}}$ de las sucesiones de números complejos. \mathcal{E} es un espacio vectorial de dimensión k ; luego asociando a cada $a = (a_1, \dots, a_k) \in \mathbf{C}^k$, la sucesión $(u_n) = f(a)$, única, que verifica las relaciones (5) y las «condiciones iniciales» $u_0 = a_1, u_1 = a_2, \dots, u_{k-1} = a_k$, se obtiene un isomorfismo f de \mathbf{C}^k en \mathcal{E} .

a) Se trata de hallar las soluciones particulares de (5) de la forma: $u_n = \lambda^n$ ($\lambda \in \mathbf{C}^*$). Puesto que $\lambda \neq 0$, (5) se escribe, después de dividir por λ^n :

$$(6) \quad \lambda^k - \alpha_{k-1} \lambda^{k-1} - \dots - \alpha_0 = 0.$$

Más adelante veremos que el primer miembro de (6) es el polinomio característico de un cierto endomorfismo de \mathcal{E} .

Supongamos que (6) posee k raíces *distintas* $\lambda_1, \lambda_2, \dots, \lambda_k$. Demostraremos que las k soluciones $(\lambda_1^n)_{n \geq 0}, (\lambda_2^n)_{n \geq 0}, \dots, (\lambda_k^n)_{n \geq 0}$ de (5) son linealmente independientes. Para ello bastará con observar que el determinante $\det [\lambda_i^{j-1}]_{1 \leq i, j \leq k}$ de sus k primeras coordenadas, igual a $\prod_{i < j} (\lambda_j - \lambda_i)$, es $\neq 0$ (fórmula de Vandermonde). Puesto que el espacio \mathcal{E} es de dimensión k , vemos que las sucesiones $(\lambda_i^n)_{n \geq 0}$ forman una base de \mathcal{E} ; luego, en este caso elemental, toda sucesión solución de (5) es de la forma:

$$(7) \quad U_n = \sum_{i=1}^k \rho_i \lambda_i^n \quad (\rho_i \in \mathbf{C});$$

las constantes ρ_i vienen determinadas por las k primeras ecuaciones (7), en donde u_0, \dots, u_{k-1} están dados de antemano. Estas ecuaciones forman un sistema de Cramer respecto de las ρ_i , cuyo determinante es precisamente $\det [\lambda_i^{j-1}]_{1 \leq i, j \leq k}$.

b) Vamos a encontrar, de nuevo, el resultado anterior, si tratamos el problema con toda generalidad.

Sea φ el endomorfismo de $\mathbf{C}^{\mathbf{N}}$ que, a cada sucesión (u_n) , le hace corresponder la sucesión (v_n) tal que: $v_n = u_{n+1}$ para todo $n \geq 0$ (traslación de $+1$ en los índices) y sea $\varphi_{\mathcal{E}}$ la restricción de φ a \mathcal{E} . $\varphi_{\mathcal{E}}$ es biyectiva, su matriz es:

$$\Phi = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \vdots & \vdots & 0 & 1 \\ \alpha_0 & \alpha_1 & \dots & \dots & \alpha_{k-1} \end{bmatrix}.$$

Sea $P(X) = \prod_{i=1}^p (\lambda_i - X)^{r_i}$ el polinomio característico de $\varphi_{\mathcal{E}}$, en donde $\lambda_1, \dots, \lambda_p$ designan los valores propios distintos. Según la teoría, el espacio \mathcal{E} es la suma directa de los subespacios N_i (N_i designa el núcleo del endomorfismo)

$$(\lambda_i \mathcal{O} - \varphi)^{r_i} \text{ de } \mathcal{E}), \text{ y sabemos que } \dim(N_i) = r_i.$$

Un cálculo directo prueba

$$(-1)^k P(X) = X^k - \alpha_{k-1} X^{k-1} - \dots - \alpha_0.$$

Resulta que toda sucesión (u_n) de $\mathbf{C}^{\mathbf{N}}$, que pertenezca al núcleo de $(\lambda_i \mathcal{O} - \varphi)^{r_i}$ verifica (5) y, por lo tanto, pertenece a \mathcal{E} . Luego N_i es también el núcleo del operador $\psi_i = (\lambda_i \mathcal{O} - \varphi)^{r_i}$ de $\mathbf{C}^{\mathbf{N}}$.

Utilicemos la identidad

$$(8) \quad \sum_{k=0}^n (-1)^k \binom{n}{k} (X - k)^p = 0, \text{ válida para } p < n.$$

(cf. ejercicio IV.18).

El núcleo de ψ_i está formado por las sucesiones (u_n) que verifican, para todo n , la relación

$$(9) \quad u_{n+r_i} - \binom{r_i}{1} \lambda_i u_{n+r_i-1} + \binom{r_i}{2} \lambda_i^2 u_{n+r_i-2} + \dots + (-1)^{r_i} \lambda_i^{r_i} u_n = 0;$$

reemplazando en (8) n por r_i , p por q y X por $n + r$, vemos que las r_i sucesiones $u_n = \lambda_i^n n^q$ ($0 \leq q \leq r_i - 1$) verifican (9). Dado que son linealmente independientes (calcular el determinante de sus r_i primeras componentes), forman una base de N_i .

Para terminar: las soluciones de (5) son las combinaciones lineales de las sucesiones del tipo $u_n = \lambda_i^n n^q$, $1 \leq i \leq p$, $0 \leq q \leq r_i - 1$.

Nota. Para $p = n$, (8) se convierte en

$$(10) \quad \sum_{k=0}^n (-1)^k \binom{n}{k} (X - k)^n = n!.$$

Deducimos que la sucesión: $u_n = \lambda_i^n n^{r_i-1}$ no pertenece al núcleo de $(\lambda_i \mathcal{D} - \varphi)^{r_i-1}$, si bien es un elemento de N_i , luego el polinomio minimal de la restricción de φ a N_i es $(X - \lambda_i)^{ir}$. Se deduce, pues, que el polinomio característico de φ_i (y de Φ) es igual (salvo el signo) a su polinomio minimal.

Ejemplo

Busquemos las sucesiones (u_n) de números complejos tales que

$$(11) \quad u_n - 2 \cos \theta \cdot u_{n-1} + u_{n-2} = 0 \quad (\theta, \text{ número complejo fijo}).$$

La ecuación característica es

$$X^2 - 2 \cos \theta \cdot X + 1 = 0.$$

Si $\cos \theta = \varepsilon$ ($\varepsilon = \pm 1$), las sucesiones solución son de la forma:

$$u_n = \varepsilon^n (a + bn) \quad (a, b \text{ constantes}).$$

En los otros casos, las sucesiones son de la forma:

$$u_n = a e^{in\theta} + b e^{-in\theta} \quad (a, b, \text{ constantes}).$$

Si u_0, u_1 están dados, la relación (11), escrita para $n = 0$ y $n = 1$, determina a y b .

Por ejemplo, si $u_0 = \cos \theta$ y $u_1 = \cos 2\theta$, la sucesión (u_n) es

$$u_n = \cos (n + 1) \theta.$$

Si $u_0 = 2 \cos \theta$ y $u_1 = 4 \cos^2 \theta - 1$, se obtendrá:

$$u_n = \frac{\sin (n + 2) \theta}{\sin \theta}.$$

En estos dos últimos casos, u_n es un polinomio en $\cos \theta$; la fórmula (11) permite calcular este polinomio directamente por recurrencia (cf. p. 160).

De forma más general, el término n -ésimo de la sucesión (u_n) definido por

$$u_n = \alpha u_{n-1} + \beta u_{n-2} \quad (n \geq 2) \quad \text{y} \quad u_1 = a_1, u_0 = a_0,$$

se puede expresar sin hacer intervenir las raíces de la ecuación característica $\lambda^2 - \alpha\lambda - \beta = 0$. En efecto, haciendo $u_n = P_n a_1 + Q_n a_0$, se tiene

$$P_1 = 1, Q_1 = 0 \quad P_2 = \alpha, Q_2 = \beta, \quad P_{n+1} = \alpha P_n + \beta P_{n-1}, \quad Q_{n+1} = \alpha Q_n + \beta Q_{n-1}.$$

Por recurrencia sobre n (utilizando la fórmula de Pascal) se comprueban las relaciones

$$P_n = \sum_{0 \leq k \leq \frac{n-1}{2}} \binom{n-1-k}{k} \alpha^{n-1-2k} \beta^k, \quad Q_n = \beta P_{n-1}.$$

§ XI.5 ENDOMORFISMOS DIAGONALIZABLES

● Supondremos en todo momento que el cuerpo de base K es algebraicamente cerrado. Hemos definido en el § 2 la notación de endomorfismo diagonalizable. La teoría del polinomio minimal permite caracterizar de forma notable tales endomorfismos:

TEOREMA XI.5.1

|| Sea u un endomorfismo del espacio vectorial E de dimensión n . Entonces u es diagonalizable si, y sólo si, su polinomio minimal sólo posee raíces simples.

Demostración

a) Si u es diagonalizable, sean $\lambda_1, \dots, \lambda_p$ sus valores propios distintos, de multiplicidades respectivas $\alpha_1, \dots, \alpha_p$. Por hipótesis, existe una base $(e_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq \alpha_i}}$ de vectores propios, siendo e_{ij} (para $j = 1, \dots, \alpha_i$) un vector propio asociado a λ_i . Cualquiera que sean los índices i, j , el vector e_{ij} pertenece al núcleo de $\lambda_i \mathcal{D} - u$, luego (puesto que los endomorfismos $\lambda_i \mathcal{D} - u$ conmutan) al núcleo de

$$r(u) = \prod_{i=1}^p (\lambda_i \mathcal{D} - u).$$

Resulta, pues, que $r(u)$ es el endomorfismo nulo. El polinomio minimal q_u es, pues, un divisor de $r(X) = \prod_{i=1}^p (\lambda_i - X)$, lo que demuestra que q_u posee únicamente raíces simples, y también (según XI.4.2) que $q_u = (-1)^p r$.

b) Si el polinomio minimal q_u de u posee únicamente raíces simples, y si $\lambda_1, \dots, \lambda_p$ son los valores propios distintos de u , se tiene:

$$q_u = \prod_{i=1}^p (X - \lambda_i) = (-1)^p P_u.$$

Según el teorema XI.3.3, E es la suma directa de los núcleos de los $\lambda_i \mathcal{O} - u$, que son precisamente los subespacios propios de u .]

Algunas aplicaciones

1) Sea $A \in M_n(\mathbf{C})$ una matriz tal que $A^m = I_n$. Entonces A es diagonalizable. En efecto su polinomio minimal es un divisor de $X^m - 1$, que posee todas sus raíces simples. Vemos además que los valores propios de A son raíces m -ésimas de 1.

2) Sea E un \mathbf{C} -espacio vectorial de dimensión finita, y sea u un endomorfismo diagonalizable de E . Designamos por F a un subespacio estable de E por u , es decir, tal que $u(F) \subset F$, y por u_F a la restricción de u a F . Para todo polinomio $M \in K[X]$, la relación

$$M(u) = 0 \quad \text{implica} \quad M(u_F) = 0.$$

Luego el polinomio minimal de u_F divide al de u . Resulta que u_F posee únicamente raíces simples, luego es diagonalizable. Los subespacios propios de u_F están contenidos en los subespacios propios de u . Se deduce fácilmente que F admite un suplementario G estable para u , que tiene además una base formada por vectores propios de u . (Esta propiedad no subsiste en el caso en que el endomorfismo no es diagonalizable.)

3) Con la ayuda del ejemplo 2) es fácil ver que si dos endomorfismos $u, v \in \mathcal{L}(E)$ conmutan ($uv = vu$) y son diagonalizables, admiten una base común de vectores propios.

En efecto, todo espacio propio V de v es estable para u (si V está asociado al valor propio λ y si $x \in V$, se tiene $v(x) = \lambda x$, de donde

$$v[u(x)] = u[v(x)] = \lambda u(x), \quad \text{luego } u(x) \in V.$$

La restricción de u a V es diagonalizable, en virtud del ejemplo 2). Si en cada V tomamos una base de vectores propios de u , y formamos la reunión de dichas bases, se obtiene una base de E cuyos elementos son a la vez vectores propios de u y vectores propios de v .

§ XI.6 ENDOMORFISMOS NILPOTENTES. FACTORES INVARIANTES. REDUCCIÓN DE JORDAN

● K designa un cuerpo algebraicamente cerrado, y E un espacio vectorial de dimensión finita n sobre K . Un endomorfismo $u \in \mathcal{L}(E)$ es **nilpotente** si existe

un entero m tal que $u^m = 0$. Cuando esto ocurre, el polinomio minimal de u divide a X^m , por lo tanto, es de la forma X^p ($1 \leq p \leq n$). El polinomio característico es $(-X)^n$, y puesto que todos los valores propios son nulos, si suponemos que $u \neq 0$ (es decir $p > 1$), entonces u nunca será diagonalizable. Entonces a las matrices asociadas a u se les llama **nilpotentes**.

Recíprocamente, si el polinomio característico de u es $(-X)^n$ (es decir, si todos los valores propios de u son nulos), u es nilpotente (Teorema de Hamilton-Cayley).

Cuando $p > 1$, el problema de la reducción se enuncia: *se trata de determinar la forma reducida de un endomorfismo nilpotente u , de polinomio minimal X^p , $p > 1$.*

1) *Caso en que $p = n$.* En este caso tenemos $u^{n-1} \neq 0$. Luego existe $e \in E$ tal que $u^{n-1}(e) \neq 0$. Hacemos $e_n = e$ y definimos la sucesión (e_i) por

$$e_{n-p} = u^p(e) = u(e_{n-p+1}) \quad (p = 1, 2, \dots, n-1).$$

Los vectores e_1, \dots, e_n obtenidos son independientes ya que $\sum_{i=1}^n \lambda_i e_i = 0$ implica

$$\sum_{i=1}^n \lambda_i u^p(e_i) = 0 = \sum_{i=p+1}^n \lambda_i e_{i-p} = 0;$$

y si en esta relación hacemos sucesivamente $p = n-1, p = n-2, \dots, p = 1$, obtenemos $\lambda_n = \lambda_{n-1} = \dots = \lambda_1 = 0$. Se obtiene pues una base (e_i) tal que $u(e_i) = e_{i-1}$ si $i > 1$ y $u(e_1) = 0$. En esta base, la matriz U de u tiene la forma siguiente, llamada *de Jordan*:

$$U = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}.$$

Recíprocamente, si U es de la forma citada, se tiene $u(e_1) = 0$, $u(e_{i+1}) = e_i$ para $i = 1, 2, \dots, n-1$, luego $u^k \neq 0$ para $k < n$, y el polinomio minimal de u es X^n .

2) *Caso general.* Se pretende descomponer u en suma directa de endomorfismos del tipo precedente.

Designamos por N_i el núcleo de u^i ($0 \leq i \leq p$). Puesto que

$$u^i(x) = 0 \quad \text{implica} \quad u^{i+1}(x) = 0,$$

se tienen las relaciones de inclusión

$$\{0\} = N_0 \subset N_1 \subset \dots \subset N_{i-1} \subset N_i \subset \dots \subset N_p = E.$$

(Veremos en seguida que todas estas inclusiones son estrictas.)

Puesto que $u^{p-1} \neq 0$, se tiene: $N_{p-1} \neq E$. Sea M_p un suplementario cualquiera de N_{p-1} en E . Vamos a definir por recurrencia descendente una sucesión $(M_k)_{1 \leq k \leq p}$ de subespacios no nulos de E , tales que:

a) N_k es la suma directa de N_{k-1} y M_k , de donde:

$$M_k \subset N_k \quad \text{y} \quad E = M_1 + M_2 + \cdots + M_p;$$

b) para cada índice k , es $u(M_k) \subset M_{k-1}$.

Supongamos contruidos M_k, M_{k+1}, \dots, M_p , y sea x un elemento *no nulo* de M_k . Se tiene $u^k(x) = 0$, luego $u(x) \in N_{k-1}$ (ya que $x \in N_k$) y

$$u(x) \notin N_{k-2} \quad (\text{ya que } x \notin N_{k-1}).$$

En primer lugar, resulta que la restricción de u a M_k es inyectiva, puesto que

$$u(M_k) \cap N_{k-2} = \{0\} \quad \text{y} \quad u(M_k) \subset N_{k-1}.$$

Existe, pues, un suplementario por lo menos de N_{k-2} en N_{k-1} , que contiene a $u(M_k)$, y entonces es suficiente tomar M_{k-1} como uno de tales suplementarios. Es claro que E es la suma directa de los M_k ($k = 1, 2, \dots, p$) y que la sucesión $m_k = \dim(M_k)$ satisface a la relación $m_k \geq m_{k+1}$. Definamos ahora por recurrencia una base de E , a saber $(\varepsilon_{i,j})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m_i}}$, tal que $(\varepsilon_{p,1}, \dots, \varepsilon_{p,m_p})$ sea una base de M_p .

Si los $(\varepsilon_{i,j})$ están contruidos para $k \leq i \leq p$, hacemos $\varepsilon_{k-1,j} = u(\varepsilon_{k,j})$ para $1 \leq j \leq m_k$ y, para $m_k < j \leq m_{k-1}$, se toma como $(\varepsilon_{k-1,j})$ una completación cualquiera de la familia libre $(\varepsilon_{k-1,j})_{1 \leq j \leq m_k}$ en una base de M_{k-1} . Reordenamos esta base, en una base $(e_{i,j})$, haciendo:

$$e_{i,j} = \varepsilon_{j,i} \quad (1 \leq i \leq m_1, 1 \leq j \leq n_i, \text{ en donde } n_i = \sup_{m_j \geq i} (j)).$$

Designamos por E_i ($1 \leq i \leq m_1$) al subespacio de E engendrado por

$$e_{i,1}, e_{i,2}, \dots, e_{i,n_i}.$$

Puesto que se tiene:

$$u(e_{i,1}) = 0, \quad u(e_{i,2}) = e_{i,1}, \dots, u(e_{i,n_i}) = e_{i,n_i-1},$$

E_i es estable para u , y la matriz U_i de u en la base $(e_{i,1}, \dots, e_{i,n_i})$ de E_i es de la forma:

$$(1) \quad U_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & \ddots & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}.$$

En la base $(e_{i,j})$ la matriz U de u es, pues, de la forma:

$$(2) \quad U = \begin{bmatrix} U_1 & 0 & \dots & 0 \\ 0 & U_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & U_{m_1} \end{bmatrix}$$

(«cuadro diagonal» de las matrices U_i : la diagonal principal de U_i está contenida en la de U).

U_i es cuadrada de orden n_i , y se observa que $n_1 = p$. El polinomio minimal de cada U_i es X^{n_i} .

DEFINICIÓN XI.6.1

$\left\{ \begin{array}{l} \text{A una matriz de la forma (1) se le llama } \mathbf{matriz \ nilpotente \ de \ Jordan.} \\ \mathbf{factores \ invariantes \ del \ endomorfismo \ } u. \end{array} \right.$

Podemos resumir las consideraciones precedentes en el siguiente:

TEOREMA XI.6.1

$\left\| \begin{array}{l} \text{Toda matriz nilpotente es semejante a un cuadro diagonal } U_1, \dots, U_{m_1} \\ \text{de matrices de Jordan de la forma (1), en donde } U_i \text{ es una matriz nilpo-} \\ \text{tente de Jordan de orden } n_i \text{ (} n_1 \geq n_2 \geq \dots \geq n_{m_1} \text{).} \end{array} \right.$

Conservamos las notaciones que preceden a la definición XI.5.1, el procedimiento empleado para construir la base (e_{ij}) prueba que *no es única*. Sin embargo, el conocimiento de la base (e_{ij}) determina los enteros n_i ($n_1 \geq n_2 \geq \dots \geq n_{m_1}$); y el conocimiento de los enteros n_i determina los enteros m_j por medio de las fórmulas: $m_j = \inf_{n_i \geq j} (i)$. El conocimiento de la base (e_{ij}) permite, pues, rehacer la sucesión de los núcleos N_k , cuya dimensión depende únicamente de u . En otras palabras, la sucesión de los enteros (n_i) , y el entero m_1 , están unívocamente determinados por u . O también: la forma (2) de una reducida de Jordan del endomorfismo nilpotente u es única.

DEFINICIÓN XI.6.2

$\left\{ \begin{array}{l} \text{Si } n_1, n_2, \dots, n_{m_1} \text{ son los enteros (no necesariamente distintos) definidos} \\ \text{anteriormente, a los polinomios minimales } X^{n_1}, X^{n_2}, \dots, X^{n_{m_1}} \text{ se les llaman} \\ \mathbf{factores \ invariantes \ del \ endomorfismo \ } u. \end{array} \right.$

Podemos, pues, decir que dos matrices nilpotentes son semejantes si, y sólo si, poseen los mismos factores invariantes. Las sucesiones de factores invariantes *clasifican*, pues, las matrices nilpotentes, salvo para la relación « A es semejante a B ».

Reducción de Jordan de una matriz cualquiera

Sea u un endomorfismo cualquiera del espacio vectorial E de dimensión n . Según los teoremas XI.3.2 y XI.3.3, si

$$P_u = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i} \quad \text{y} \quad q_u = \prod_{i=1}^p (\lambda_i - X)^{\beta_i}$$

son sus polinomios característico y minimal, el espacio característico $N_i = \text{Ker} (\lambda_i \mathcal{O} - u)^{\beta_i}$ es estable, de dimensión α_i , y E es suma directa de los N_i . Además, si u_i designa la restricción de u a N_i , los polinomios característico y minimal de u_i son $(\lambda_i - X)^{\alpha_i}$ y $(\lambda_i - X)^{\beta_i}$. Para obtener una forma reducida de u , es suficiente hallar una forma reducida de cada uno de los u_i . Ponemos:

$$(3) \quad u_i = \lambda_i \mathcal{O}_i + (u_i - \lambda_i \mathcal{O}_i) = \lambda_i \mathcal{O}_i + v_i \quad (\mathcal{O}_i : \text{aplicación idéntica de } N_i).$$

v_i es nilpotente, de polinomio minimal X^{β_i} . Así u_i es la suma de una homotecia y de un endomorfismo nilpotente. Además (dado que λ_i es el único valor propio de u_i) la descomposición de (3) es única. En toda base de N_i , la matriz de $\lambda_i \mathcal{O}$ es $\lambda_i I_{\alpha_i}$. Además, si $X^{n_{i,1}}, X^{n_{i,2}}, \dots, X^{n_{i,r_i}}$ son los factores invariantes de v_i , existe una base B_i de N_i en que la matriz de v_i es el cuadro diagonal de las matrices

$$U_{i,k} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 & \vdots \\ 0 & \dots & \dots & 0 & 0 \end{bmatrix} \quad (1 \leq k \leq r_i) \text{ cuadradas de orden } n_{i,k} \quad (n_{i,1} = \beta_i).$$

En la base B_i , la matriz J_i de u_i es, pues, el cuadro diagonal de las matrices

$$J_{i,k} = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 & \vdots \\ 0 & \dots & \dots & 0 & \lambda_i \end{bmatrix}, \quad 1 \leq k \leq r_i, \quad J_{i,k} \text{ cuadrada de orden } n_{i,k}.$$

En la base $B = \bigcup_{1 \leq i \leq p} B_i$, la matriz J de u es, entonces, el cuadro diagonal de las matrices $J_{i,k}$ ($1 \leq i \leq p$, $1 \leq k \leq r_i$). El aspecto de J es el siguiente:

$$(4) \quad M = \begin{bmatrix} \boxed{M_1} & 0 & \dots & 0 \\ 0 & \boxed{M_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \boxed{M_p} \end{bmatrix}$$

M_i es cuadrada de orden α_i , es el cuadro diagonal de las $J_{i,k}$ ($1 \leq k \leq r_i$). En la diagonal de M_i figura el valor propio λ_i .

$$M_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & & & \\ \vdots & & \ddots & & \\ 0 & \dots & & \lambda_i & \end{bmatrix}.$$

Encima de los términos λ_i figuran 1 ó 0. En los demás lugares figuran 0.

El número máximo de términos «1» consecutivos es $\beta_i - 1$.

A la forma (4) de la matriz de u en la base B se la llama *forma reducida de Jordan*. Si cada β_i es igual a 1, la matriz M es diagonal, lo que está de acuerdo con el teorema XI.5.1.

En el transcurso de este estudio, hemos observado que cada u_i es suma de una homotecia y de un endomorfismo nilpotente. En general, se tiene:

TEOREMA XI.6.2

Para todo endomorfismo $u \in \mathcal{L}_K(E)$, existe un único par (Δ, v) de endomorfismos de E tal que:

- a) Δ es diagonalizable y v es nilpotente.
- b) $u = \Delta + v$.
- c) $v\Delta = \Delta v$.

Demostración (abreviada)

1) La existencia resulta fácilmente del estudio que precede. Conservando las notaciones de este estudio, es suficiente tomar, evidentemente:

$$\begin{aligned} \Delta &= \text{suma directa de las homotecias } \lambda_i \mathcal{O}_i \\ v &= u - \Delta. \end{aligned}$$

2) *Unicidad*: de $v\Delta = \Delta v$, se deduce que todo subespacio propio de Δ es v -estable (pues $\Delta(x) = \lambda x$ implica $\Delta(v(x)) = v(\Delta(x)) = v(\lambda x) = \lambda \cdot v(x)$). Triangulando la restricción de v a cada subespacio propio de Δ , se deduce que los polinomios característicos de Δ y $\Delta + v$ son iguales, ya que los subespacios propios de Δ son también los subespacios característicos de $u = \Delta + v$. Esto implica que la restricción de Δ a N_i es $\lambda_i \mathcal{O}_i$. c.q.d.

Resultado práctico acerca de la reducción de Jordan

Volvemos al problema general de la reducción de Jordan que hemos analizado. Una consecuencia importante de nuestro estudio lo constituye el resultado XI.6.3 que sigue, y que constituye la forma «utilizable» de la teoría de las reducidas de Jordan. A fin de enunciar cómodamente este resultado, es conveniente llamar **matriz de Jordan** a toda matriz cuadrada de orden 1 (o sea $J = [\lambda]$) y a toda matriz cuadrada de orden ≥ 2 de la forma:

$$J = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & \lambda \end{bmatrix};$$

en ambos casos λ designa un elemento de K , que es el único valor propio de J . Entonces se tiene:

XI.6.3 Sea u un endomorfismo de un K -espacio vectorial E de dimensión finita, en donde K es un cuerpo algebraicamente cerrado. Existe una base $(e_i)_{1 \leq i \leq n}$ de E en que la matriz M de u es un **cuadro diagonal de matrices de Jordan**, es decir, de la forma

$$(5) \quad M = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_s \end{bmatrix}$$

en donde J_1, J_2, \dots, J_s designan matrices de Jordan.

Se observará que el conjunto de los valores propios de las diversas matrices J_k es igual al conjunto de los valores propios de M ; pero debe tenerse en cuenta que, si $p \neq q$, los valores propios de J_p y J_q no son necesariamente distintos.

Observemos, finalmente, que la matriz M definida por (5) es diagonalizable si, y sólo si, cada J_k es de orden 1.

Capítulo XII

Formas bilineales y formas cuadráticas

§ XII.1 GENERALIDADES ACERCA DE LAS FORMAS BILINEALES

- En lo que sigue, el cuerpo base K lo supondremos conmutativo, de característica $\neq 2$.

Recordemos una definición dada ya en el capítulo X.

DEFINICIÓN XII.1.1

Una forma bilineal sobre el K -espacio vectorial E es una aplicación $B : E \times E \rightarrow K$, lineal con respecto a cada una de las componentes, es decir, que verifique

$$(1) \quad B(\lambda x, y) = B(x, \lambda y) = \lambda B(x, y)$$

para todo $x \in E$, todo $y \in E$ y todo $\lambda \in E$,

$$B(x + y, z) = B(x, z) + B(y, z), \quad B(x, y + z) = B(x, y) + B(x, z)$$

para todo $x \in E$, todo $y \in E$, todo $z \in E$.

Si E es de dimensión finita n , y si (e_1, \dots, e_n) es una base de E , B queda unívocamente determinada dando los n^2 escalares $b_{ij} = B(e_i, e_j)$ ($i, j = 1, 2, \dots, n$), en virtud de la fórmula, deducida de (1)

$$(2) \quad B\left(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^n \mu_j e_j\right) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j b_{ij}.$$

Con mayor precisión, la relación (2) prueba que el espacio $B(E)$ de las formas bilineales sobre E está engendrado por las n^2 formas B_{ij} definidas por

$$(3) \quad B_{ij}(e_i, e_j) = 1 \quad \text{y} \quad B_{ij}(e_k, e_l) = 0 \quad \text{si} \quad (k, l) \neq (i, j).$$

Además estas formas son linealmente independientes, pues una relación del tipo $\sum_{i,j} \lambda_{ij} B_{ij} = 0$ implica $\sum_{i,j} \lambda_{ij} B_{ij}(e_k, e_l) = 0$, es decir, $\lambda_{kl} = 0$, cualesquiera que sean $k, l = 1, 2, \dots, n$.

Por lo tanto, podemos enunciar:

XII.1.1 *El espacio vectorial $B(E)$ de las formas bilineales sobre un espacio vectorial E , de dimensión finita n , tiene dimensión n^2 . Y a cada base $(e_i)_{1 \leq i \leq n}$ de E corresponde la base de $B(E)$ formada por las n^2 formas B_{ij} definidas por (3).*

A lo largo de este capítulo nos ocuparemos esencialmente de las formas bilineales sobre un espacio vectorial de dimensión finita.

Matriz asociada

Resulta de XII.1.1 que $B(E)$ es isomorfo al espacio de las matrices cuadradas de orden n con elementos de K . A la matriz $M = [b_{ij}]$ definida por $b_{ij} = B(e_i, e_j)$ se le denominará *matriz de B en la base (e_1, \dots, e_n)* .

Cambio de base

Sean (e_1, \dots, e_n) y (f_1, \dots, f_n) dos bases del espacio vectorial E , y sea P la matriz que pasa de la primera a la segunda. Consideraremos, por otra parte, una forma bilineal B sobre E . Sea $M = [B(e_i, e_j)]$ la matriz de B en la base (e_1, \dots, e_n) y $N = [B(f_i, f_j)]$ la de B en la base (f_1, \dots, f_n) . Vamos a buscar la relación existente entre M y N .

Para ello observemos que (2) se puede escribir en forma matricial. Si $x = \sum x_i e_i$

e $y = \sum y_i e_i$, y designamos por \mathcal{X} y por \mathcal{Y} a las matrices columna $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, entonces (2) se escribe de la manera siguiente:

$$(4) \quad B(x, y) = {}^t\mathcal{X} M \mathcal{Y} = {}^t\mathcal{Y} {}^tM \mathcal{X},$$

● y la matriz M queda totalmente determinada por una u otra de estas identidades (dada la forma B).

Análogamente, si $x = \sum x'_i f_i$ e $y = \sum y'_i f_i$, se tiene:

$$(5) \quad B(x, y) = {}^t\mathcal{X}' N \mathcal{Y}' = {}^t\mathcal{Y}' {}^tN \mathcal{X}'.$$

Finalmente sabemos que $\mathcal{X} = P\mathcal{X}'$, $\mathcal{Y} = P\mathcal{Y}'$, de donde ${}^t\mathcal{X} = {}^t\mathcal{X}' {}^tP$ y ${}^t\mathcal{Y} = {}^t\mathcal{Y}' {}^tP$. Combinándolo con (4) y (5), obtenemos:

$$B(x, y) = {}^t\mathcal{X}' {}^tPMP\mathcal{Y}' = {}^t\mathcal{X}' N \mathcal{Y}',$$

de donde deducimos, puesto que N está caracterizada por (5),

$$(6) \quad \boxed{N = {}^tPMP}.$$

Puesto que P es regular, el rango de N es igual al de M . Este rango depende, pues, únicamente de B . Vamos a precisar este resultado:

Aplicaciones I y J

Sea B una forma bilineal sobre el K -espacio E . Para todo $y \in E$, definimos la forma lineal $B(\cdot, y)$ por $(B(\cdot, y))(x) = B(x, y)$. (La forma $B(\cdot, y)$ se designa también por B^y .) Análogamente sea $B(x, \cdot)$ la forma lineal $y \mapsto B(x, y)$ sobre E . (La forma $B(x, \cdot)$ se designa también a veces por B_x .)

La aplicación $I : y \mapsto B(\cdot, y)$ es lineal de E en el dual E^* .

La aplicación $J : x \mapsto B(x, \cdot)$ es lineal de E en el dual E^* .

Suponemos E de dimensión finita: entonces es posible identificar E y E^{**} , y vemos que las aplicaciones I y J son traspuestas una de otra, por lo que tienen el mismo rango (T. VIII.5.7).

DEFINICIÓN XII.1.2

§ Sea E un espacio vectorial de dimensión finita y B una forma bilineal sobre E . Al rango común a las aplicaciones I y J , anteriormente definidas se le llama **rango de B** .

TEOREMA XII.1.2

|| El rango de una forma bilineal sobre un espacio vectorial de dimensión finita E es igual al rango de la matriz de esta forma en cualquier base de E .

Demostración. Si (e_1, \dots, e_n) designa una base de E y $M = [b_{ij}]$ la matriz de la forma bilineal B en esta base, y (e_1^*, \dots, e_n^*) la base dual de (e_1, \dots, e_n) , se verifica

$$I(e_i) = \sum_j b_{ji} e_j^*,$$

de forma que M es también la matriz de I en las bases (e_i) y (e_i^*) , de donde se sigue el resultado deseado. c.q.d.

● El teorema XII.1.2 proporciona un método práctico para calcular el rango de una forma bilineal. Cuando este rango es máximo, se dice que la forma es **no degenerada**; en caso contrario se le llama **degenerada**. Una forma no degenerada se halla, pues, caracterizada por el hecho de que el determinante de su matriz en una base cualquiera —llamado **discriminante** de la forma en esta base— es distinto de cero.

Nota. Si E no tiene dimensión finita, se dice que B es *no degenerada* si las dos aplicaciones I y J son inyectivas.

§ XII.2 FORMAS BILINEALES SIMÉTRICAS Y FORMAS CUADRÁTICAS

La idea básica la constituye ahora el estudio de los polinomios homogéneos de grado dos. Si B es una forma bilineal sobre el espacio E , de dimensión n , la aplicación $x \mapsto B(x, x)$ es un polinomio homogéneo de grado 2 respecto de las coordenadas de x en cualquier base (e_1, \dots, e_n) ; pues, si hacemos: $x = \sum x_i e_i$, la fórmula (2) del § 1 nos da:

$$(1) \quad B(x, x) = \sum_{i=1}^n \sum_{j=1}^n b_{ij} x_i x_j, \quad \text{con} \quad b_{ij} = B(e_i, e_j).$$

Sin embargo, el *mismo* polinomio de grado 2 se puede obtener a partir de formas bilineales *distintas*. Pero si nos limitamos a las formas bilineales *simétricas*, esta indeterminación desaparece:

DEFINICIÓN XII.2.1

Una forma bilineal B sobre el espacio vectorial E es **simétrica** si cumple:

$$B(x, y) = B(y, x) \quad \text{para todo } x \in E \text{ y todo } y \in E.$$

Caso en que E es de dimensión finita

Es inmediato que *la forma B es simétrica si, y sólo si, su matriz en cualquier base de E es una matriz simétrica*, en virtud de la misma definición de esta matriz (cf. fórmula (2) del § 1).

Esto muestra que *el espacio vectorial $\mathcal{S}_2(E)$ de las formas bilineales simétricas sobre E es isomorfo al espacio vectorial de las matrices cuadradas de orden n simétricas con elementos en K* . (No existe un isomorfismo canónico, pues un isomorfismo entre estos dos espacios vectoriales depende de la base dada en E .) $\mathcal{S}_2(E)$ tiene, por lo tanto, dimensión $n(n+1)/2$.

Sea B una forma bilineal simétrica sobre E , y pongamos $Q(x) = B(x, x)$. Las fórmulas (1) y (2) del § 1 nos dan las siguientes propiedades:

$$(2) \quad \begin{cases} Q(\lambda x) = \lambda^2 Q(x) & \text{para todo } x \in E, \text{ y todo } \lambda \in K; \\ B(x, y) = \frac{1}{2} [Q(x+y) - Q(x) - Q(y)] & \text{para todo } x \in E \text{ y todo } y \in E. \end{cases}$$

Estas fórmulas nos permiten identificar el espacio $\mathcal{S}_2(E)$ con el espacio de las formas cuadráticas sobre E :

DEFINICIÓN XII.2.2

Sea E un espacio vectorial de dimensión finita. Una **forma cuadrática** sobre E es una aplicación $Q : E \rightarrow K$ que se expresa, en cada base de E , en forma de polinomio homogéneo de grado 2 en las variables coordinadas, o que es idénticamente nulo.

Puesto que las fórmulas de cambio de coordenadas son lineales es suficiente que esta condición se verifique en *una* base dada, para que se verifique en *toda* base. Quede claro que el polinomio que representa a B depende de la base elegida.

Si (e_1, \dots, e_n) designa una base de E , y si $x = \sum_i x_i e_i$, tenemos entonces

$$Q(x) = \sum_{i=1}^n a_i x_i^2 + 2 \sum_{i < j} a_{ij} x_i x_j,$$

en donde los a_{ij} son elementos fijos cualesquiera de K .

Obsérvese que los coeficientes a_i y a_{ij} están unívocamente determinados por la forma cuadrática Q (cf. (2) y por la definición de la matriz de Q en la base (e_i)).

En otras palabras (con la única condición de que K no tenga característica 2), la función polinomio

$$x \mapsto \sum_{i=1}^n a_i x_i^2 + 2 \sum_{i < j} a_{ij} x_i x_j$$

está asociada a un único *polinomio formal*, a saber:

$$\sum_{i=1}^n a_i X_i^2 + 2 \sum_{i < j} a_{ij} X_i X_j.$$

En la práctica, para dar una forma cuadrática sobre E , bastará con dar un polinomio homogéneo, nulo o de grado 2, de las coordenadas relativas a una base fija.

TEOREMA XII.2.1

Sea E un espacio vectorial de dimensión finita, y sea $Q : E \rightarrow K$ una aplicación.

— Q es una forma cuadrática si, y sólo si,

1.º $Q(\lambda x) = \lambda^2 Q(x)$ para todo $x \in E$ y todo $\lambda \in K$.

2.º La aplicación $(x, y) \mapsto B(x, y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y))$ es bilineal sobre E .

— Además, se tiene entonces: $B(x, x) = Q(x)$.

Demostración

— Si Q es cuadrática, la condición 1.º se verifica evidentemente. A fin de comprobar 2.º tomemos una base (e_1, \dots, e_n) de E y hagamos

$$x = \sum x_i e_i, \quad y = \sum y_j e_j$$

se tiene (puesto que el cuerpo de base K se ha supuesto de característica $\neq 2$):

$$Q(x) = \sum_{i=1}^n a_i x_i^2 + 2 \sum_{i < j} a_{ij} x_i x_j,$$

$$Q(x+y) = \sum_{i=1}^n a_i (x_i + y_i)^2 + 2 \sum_{i < j} a_{ij} (x_i + y_i) (x_j + y_j)$$

$$\begin{aligned} \frac{1}{2} [Q(x+y) - Q(x) - Q(y)] &= \sum_{i=1}^n a_i x_i y_i + \sum_{i < j} a_{ij} (x_i y_j + x_j y_i) \\ &= \sum_{1 \leq i \leq n, 1 \leq j \leq n} b_{ij} x_i y_j, \end{aligned}$$

haciendo $b_{ii} = a_i$ y $b_{ji} = b_{ij} = a_{ij}$ si $i < j$.

Esto demuestra (cf. fórmula (2) del § 1) que $B(x, y)$ es bilineal y simétrica. Además es claro, según la condición 2.º, que $B(x, y) = B(y, x)$.

— Recíprocamente, si se verifican 1.º y 2.º, podemos escribir:

$$B(x, x) = \frac{1}{2} (Q(2x) - 2Q(x)) = \frac{1}{2} (4Q(x) - 2Q(x)) = Q(x).$$

Puesto que B es bilineal, $x \mapsto B(x, x)$ es una forma cuadrática (cf. el principio de este §), por lo tanto Q es cuadrática. c.q.d.

Formas asociadas

A toda forma bilineal simétrica B sobre E le hacemos corresponder la forma cuadrática (llamada *asociada a B*), $B(x, x) = Q(x)$. Acabamos de ver que esta aplicación es epiyectiva, en virtud del teorema XII.2.1, 2.º. Pero también es inyectiva, en virtud de la fórmula:

$$B(x, y) = \frac{1}{2} [Q(x + y) - Q(x) - Q(y)],$$

que nos muestra que B está determinada por Q . Podemos resumir todo esto en el teorema siguiente:

TEOREMA XII.2.2

|| La aplicación que, a cada forma bilineal simétrica B sobre E , le hace corresponder su forma cuadrática asociada Q , es un isomorfismo del espacio vectorial de las formas bilineales simétricas sobre E , en el espacio de las formas cuadráticas sobre E .

Según hemos dicho antes, a Q se le llama *forma cuadrática asociada a B* ; a B se le llama *forma bilineal asociada a Q* , o la *forma polar asociada a Q* .

En virtud del teorema XII.2.2 toda noción relacionada con B se puede considerar como una noción relacionada con Q , y viceversa. Por ejemplo, se llamará *rango* de la forma cuadrática Q , al rango de su forma polar. O bien, el *discriminante* de la forma cuadrática Q en una base, será el discriminante en esta misma base de su forma polar, etc.

Para terminar este §, vamos a dar una interpretación funcional de la matriz de una forma cuadrática Q en una base (es decir, la matriz de la forma polar B de Q).

Sea (e_i) una base prefijada, con relación a ella tenemos

$$B(x, y) = \sum_{i,j} b_{ij} x_i x_j,$$

de donde

$$Q(x) = B(x, x) = \sum_{i=1}^n b_{ii} x_i^2 + 2 \sum_{i < j} b_{ij} x_i x_j.$$

Las n derivadas parciales formales:

$$\frac{\partial Q}{\partial x_i}(x) = 2 b_{ii} x_i + 2 \sum_{j > i} b_{ij} x_j + 2 \sum_{j < i} b_{ij} x_j = 2 \sum_{j=1}^n b_{ij} x_j,$$

son formas lineales sobre E , cuya matriz es $2M$ (si M designa la matriz asociada a B). Enunciaremos:

XII.2.3 La matriz M asociada a una forma cuadrática Q en una base $(e_i)_{1 \leq i \leq n}$ es la matriz de las n formas lineales Q_i definidas por

$$Q_i(x) = \frac{1}{2} \frac{\partial Q}{\partial x_i}(x) \quad 1 \leq i \leq n.$$

De donde resulta que la forma polar B de Q es

$$(3) \quad B(x, y) = \frac{1}{2} \sum_{i=1}^n Q_i(x) y_i = \frac{1}{2} \sum_{i=1}^n Q_i(y) x_i,$$

y en particular se tiene

$$(4) \quad Q(x) = \frac{1}{2} \sum_{i=1}^n Q_i(x) x_i.$$

Se observará que (4) es un caso particular de la *identidad de Euler*: todo polinomio f homogéneo de grado n verifica la identidad

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(x) = n f(x).$$

Extensión. El espacio vectorial E no es ahora necesariamente de dimensión finita. Se llama *forma cuadrática* sobre E a una aplicación Q de la forma $x \mapsto B(x, x)$, en donde B es una forma bilineal simétrica sobre E . Las aplicaciones B , Q , verifican todavía las relaciones (2); y a la forma B , que está unívocamente determinada por Q , se le llama *polar* de Q . Para $K = \mathbf{R}$, la teoría de las diferenciales (cf. tomo 2, Análisis) nos permitirá extender la fórmula (3) en la forma:

$$B(x, y) = \frac{1}{2} Q'(x) \cdot y = \frac{1}{2} Q'(y) \cdot x,$$

en donde $Q'(x)$ designa la *diferencial* de la aplicación Q en el punto x .

DEFINICIÓN XII.3.2

$\} \text{ El núcleo de la forma cuadrática } Q \text{ sobre } E \text{ es el núcleo de la aplicación}$
 $\} J : E \rightarrow E^* \text{ asociada a la forma polar } B \text{ de } Q.$

Por definición, el núcleo N de Q es, por lo tanto, el conjunto de los $x \in E$ tales que, para todo $y \in E$, se verifica $B(x, y) = 0$. Es decir: $N = E^\perp$. Se tiene entonces la proposición inmediata siguiente:

XII.3.1 Una forma cuadrática es no degenerada si, y sólo si, su núcleo se reduce a $\{0\}$.

Cuando E es de dimensión finita, para toda forma no degenerada, J es un isomorfismo de E en E^* . Enunciaremos este resultado en la forma utilizada en la práctica:

TEOREMA XII.3.2

$\| \text{ Sea } B \text{ una forma bilineal asociada a una forma cuadrática no degenerada sobre un espacio vectorial de dimensión finita. Para toda forma lineal } \varphi \text{ sobre } E \text{ existe un elemento único } y \text{ de } E \text{ tal que } \varphi(x) = B(x, y) \text{ para todo } x \in E.$

Dar una forma B de este tipo permite **identificar** E^* con E por medio del isomorfismo J , y la relación de ortogonalidad definida anteriormente equivale a la definida en el § VIII.5.

Las formas no degeneradas tienen importancia en virtud del siguiente:

TEOREMA XII.3.3

$\| \text{ Sea } Q \text{ una forma cuadrática no degenerada sobre el espacio vectorial } E \text{ (de dimensión finita } n \text{)}. \text{ Para todo subespacio } H \text{ de } E, \text{ se tiene}$

$$(1) \quad \dim(H) + \dim(H^\perp) = n,$$

$$(2) \quad (H^\perp)^\perp = H.$$

$\| \text{ De ahí que, si } A \text{ es una parte cualquiera de } E: (A^\perp)^\perp = \text{Vect}(A).$

Cuando B es una forma bilineal *cualquiera* sobre E y K es, de nuevo, cualquiera, se obtiene todavía una forma cuadrática $Q = \varphi(B)$ haciendo $Q(x) = B(x, x)$. La forma polar B_0 de Q está dada entonces por

$$B_0(x, y) = \frac{1}{2}[B(x, y) + B(y, x)] ;$$

y el núcleo de la aplicación φ es el espacio vectorial de las formas bilineales antisimétricas sobre E (puesto que K no tiene característica 2).

§ XII.3 ORTOGONALIDAD

DEFINICIÓN XII.3.1

Si Q designa una forma cuadrática sobre el espacio vectorial E , y B designa la forma polar asociada, a dos elementos x, y de E se les llama **ortogonales** (o **conjugados**) **respecto de Q** si se verifica:
 $B(x, y) = 0$ (notación: $x \perp y$ cuando no hay peligro de confusión).
 Para toda parte A de E , el **ortogonal a A** es la parte de E , designada por A^\perp , definida por la relación:

$$A^\perp = \{ y \mid y \in E \text{ y } (\forall x \in A) B(x, y) = 0 \} .$$

En otras palabras, A^\perp es el conjunto de los elementos de E ortogonales a todos los elementos de A .

Observemos que para toda parte A de E , A^\perp es un subespacio vectorial de E , y que se verifican las relaciones evidentes que siguen:

$$\{ 0 \}^\perp = E, \quad (A \cup B)^\perp = A^\perp \cap B^\perp, \quad (A \cap B)^\perp \supset A^\perp + B^\perp,$$

$$A^\perp = [\text{Vect}(A)]^\perp, \quad (A^\perp)^\perp \supset A \quad (1).$$

Si E es de dimensión finita n , y si H es un subespacio de E , de dimensión p , elijamos una base $A = (a_1, \dots, a_p)$ de H . Entonces $H^\perp = A^\perp$ es el conjunto de puntos de E que verifican las p ecuaciones lineales $B(a_i, y) = 0$ ($1 \leq i \leq p$). Se tiene entonces $\dim H^\perp \geq n - p$ (cf. Teorema X.5.2), es decir, la desigualdad:

$$(0) \quad \dim(H) + \dim(H^\perp) \geq \dim(E).$$

Núcleo

Nota preliminar. Si B es una forma bilineal *simétrica* sobre E , las aplicaciones I y J definidas en el § 1 coinciden. Esta aplicación única, $E \rightarrow E^*$, definida por $x \mapsto B(x, \cdot)$ la designaremos por J .

(1) Se recuerda que $\text{Vect}(A)$ designa el subespacio vectorial engendrado por A .

Demostración. Consideremos el isomorfismo $J : E \rightarrow E^*$ asociado a Q , y busquemos la imagen $J(H^\perp)$, que es el conjunto formado por las formas $B(., y)$ tales que $y \in H^\perp$. Con otras palabras, según XII.3.2, es el conjunto de las formas lineales φ sobre E que verifican $\varphi(x) = 0$, luego es *el ortogonal* H^0 a H en el sentido de la dualidad entre E y E^* (este razonamiento no hace más que precisar la identificación anunciada anteriormente, de las dos nociones de ortogonalidad). La relación (1) se sigue entonces del teorema VIII.5.4.

Se deduce fácilmente la relación (2), pues sabemos que $(H^\perp)^\perp \supset H$, y en virtud de (1) se tiene:

$$\dim(H) = \dim(H^\perp)^\perp \text{ . c.q.d.}$$

Nota importante. A pesar de la relación (1), en general E no es la suma directa de H y H^\perp .

Contraejemplos

- 1) $E = \mathbf{R}^2$, $B(x, y) = x_1 y_1 - x_2 y_2$; $H = \{x \mid x_1 = x_2\}$. Entonces $H^\perp = H$.
- 2) $E = \mathbf{R}^4$, $B(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4$.
 $H = \{x \mid x_1 = x_2, x_3 = x_4\}$, $H^\perp = \{y \mid y_1 + y_2 = 0 \text{ y } y_3 = y_4\}$.
 $H \cap H^\perp$ es la recta $x_1 = x_2 = 0, x_3 = x_4$.
- 3) $E = \mathbf{R}^3$, $B(x, y) = x_1 y_1 + x_2 y_2 - x_3 y_3$.
 $H = \{x \mid x_1 = x_3 \text{ y } x_2 = 0\}$, $H^\perp = \{x \mid x_1 = x_3\}$; de donde $H \subset H^\perp$.
 $\dim(H) = 1$ y $\dim(H^\perp) = 2$.

DEFINICIÓN XII.3.3

Sea E un espacio dotado de una forma cuadrática Q . Un elemento $x \in E$ es **isótropo** si $B(x, x) = 0$. Un subespacio H de E es **isótropo** si $H \cap H^\perp \neq \{0\}$, y **totalmente isótropo** si $H \subset H^\perp$ (el contraejemplo 3), dado antes, muestra la necesidad de distinguir ambas nociones).

● Si H es no isótropo, y E es de dimensión finita, la desigualdad (0) se transforma en igualdad (incluso si Q es degenerado) y prueba que E es *suma directa* de H y de H^\perp .

Si B es la forma polar de Q , decir que H es isótropo equivale a decir que la restricción de B a $H \times H$ es degenerada. Decir que H es totalmente isótropo significa que la restricción de B a $H \times H$ es nula.

Sea x un elemento *no nulo* de E , y designemos por Kx al subespacio engendrado por x . Se tiene entonces:

$$(Kx \text{ isótropo}) \Leftrightarrow (Kx \text{ totalmente isótropo}) \Leftrightarrow (Q(x) = 0).$$

DEFINICIÓN XII.3.4

Sea Q una forma cuadrática sobre el espacio vectorial de dimensión finita E . La base (e_1, \dots, e_n) de E se llama **ortogonal** (para Q) si se verifica;

$$B(e_i, e_j) = 0 \quad \text{para } i \neq j.$$

Se dice que es **ortonormal** si

$$B(e_i, e_j) = \delta_{ij} \quad (\text{símbolo de Kronecker})$$

cualesquiera que sean $i, j = 1, 2, \dots, n$.

TEOREMA XII.3.4

|| Para toda forma cuadrática Q sobre un espacio vectorial de dimensión finita E , existe una base ortogonal.

(Nota. Obsérvese que no existe necesariamente una base ortonormal, ver § 4.)

Demostración. Por recurrencia sobre $n = \dim(E)$.

— Para $n = 1$, toda base es ortogonal.

— Supongamos el teorema cierto para el entero n ; demostrémoslo para el entero $n + 1$.

Si $Q = 0$, es evidente. Si no, existe $e_{n+1} \in E$ tal que $Q(e_{n+1}) \neq 0$. Sea F el subespacio de dimensión 1 engendrado por e_{n+1} . El ortogonal F^\perp a F es el conjunto de los $y \in E$ tales que $B(e_{n+1}, y) = 0$. Es, por lo tanto, el núcleo de la forma lineal no nula $B(e_{n+1}, \cdot)$ sobre el espacio E de dimensión $n + 1$. Resulta, pues, que F^\perp tiene dimensión n , y puesto que $Q(e_{n+1}) \neq 0$, no puede ser que $F \subset F^\perp$. Luego F no es isótropo, y se tiene $F \cap F^\perp = \{0\}$. Luego E es suma directa de F y F^\perp .

Pero, en virtud de la hipótesis de recurrencia, existe una base (e_1, \dots, e_n) de F^\perp , ortogonal para la restricción de Q a F^\perp (cuya forma polar es la restricción de B a $F^\perp \times F^\perp$).

Evidentemente que $(e_1, \dots, e_n, e_{n+1})$ es una base de E , ortogonal para Q . ||

Nota. En una base (e_1, \dots, e_n) de E , ortogonal para la forma cuadrática Q , la matriz de Q es diagonal, y recíprocamente. Sean $(a_{11}, a_{22}, \dots, a_{nn})$ los elementos diagonales de esta matriz, y (x_1, \dots, x_n) las coordenadas de $x \in E$. Se tiene:

$$(3) \quad Q(x) = \sum_i a_{ii} x_i^2.$$

Recíprocamente, si la expresión de Q en la base (e_i) tiene la forma (3), esta base es ortogonal. Enunciaremos:

XII.3.5 *Para toda forma cuadrática Q sobre E existe una base de E en que la matriz de Q es diagonal y, por lo tanto, respecto a ella, se verifica*

$$Q(x) = \sum_{i=1}^n \lambda_i x_i^2 \quad \text{para cierto sistema de escalares } (\lambda_i).$$

Observemos que, en una base ortogonal (e_1, \dots, e_n) para Q , el rango de Q es evidente, pues es el número de elementos no nulos entre los elementos diagonales (λ_i) de la matriz de Q . Este número es, pues, independiente de la base ortogonal elegida, y es una constante vinculada a la forma cuadrática.

El teorema XII.3.5 es una primera aproximación del problema de la clasificación de las formas cuadráticas, que estudiaremos en el § siguiente.

§ XII.4 PROBLEMA DE LA CLASIFICACIÓN. SOLUCIÓN CUANDO $K = \mathbf{C}$ O $K = \mathbf{R}$

DEFINICIÓN XII.4.1

Dos formas cuadráticas Q_1 y Q_2 sobre el mismo espacio vectorial E son equivalentes si existe un automorfismo φ de E tal que

$$Q_2(\varphi(x)) = Q_1(x) \quad \text{para todo } x \in E.$$

Equivale a decir, si B_1 y B_2 son las formas polares de Q_1 y Q_2 , es :

$$B_2(\varphi(x), \varphi(y)) = B_1(x, y) \quad \text{para todo } x \in E \text{ y todo } y \in E.$$

La relación « Q_1 y Q_2 son equivalentes» es una relación de equivalencia sobre el conjunto de las formas cuadráticas sobre E . Si E tiene dimensión finita, la definición XII.4.1 significa simplemente que existen dos bases

$$(e_1, \dots, e_n), (f_1, \dots, f_n) \quad \text{de } E,$$

tales que la matriz de Q_1 en (e_i) es la misma que la matriz de Q_2 en (f_i) .

Una vez establecido lo anterior, el problema de la clasificación de formas cuadráticas consiste en determinar las diversas clases de equivalencia de formas cuadráticas o, si se prefiere, en buscar condiciones necesarias y suficientes para que dos formas cuadráticas sean equivalentes. Una manera de afrontar este problema consiste en buscar bases para las que la matriz de una forma sea lo más simple posible —tal como se ha hecho, en el capítulo XI, para buscar las clases de aplicaciones lineales equivalentes.

Pero como este problema, en el caso general, es difícil (por ejemplo, para cuerpos relativamente simples como el de los racionales, es preciso desarrollar toda

una teoría). Vamos a resolverlo *para los cuerpos* \mathbf{R} y \mathbf{C} . El lector encontrará, en el problema n.º 8, una aproximación del problema en el caso de *cuerpos finitos*.

Nota. El problema de la clasificación se puede enunciar utilizando el lenguaje de los grupos que operan sobre un conjunto.

$\mathcal{S}_n(K)$ designa el conjunto de las matrices cuadradas simétricas; hacemos operar el grupo $\text{GL}(n, K)$ sobre $\mathcal{S}_n(K)$ *por la derecha* estableciendo, para $P \in \text{GL}(n, K)$ y $A \in \mathcal{S}_n(K)$,

$$A * P = {}^t P A P.$$

Entonces el problema de la clasificación de formas cuadráticas consiste en caracterizar las órbitas de $\mathcal{S}_n(K)$ según $\text{GL}(n, K)$.

TEOREMA XII.4.1

Toda forma cuadrática de rango r sobre \mathbf{C}^n es equivalente a la forma

$$x_1^2 + x_2^2 + \cdots + x_r^2.$$

(En consecuencia, para que dos formas cuadráticas sobre un \mathbf{C} -espacio vectorial de dimensión n sean *equivalentes*, es necesario y suficiente que tengan el *mismo rango*.)

Demostración. Sea Q una forma de rango r sobre \mathbf{C}^n . Designamos por (f_1, \dots, f_n) una base ortogonal para Q (T. XII.3.3). Si (y_1, \dots, y_n) son las coordenadas de $x \in E$ en esta base, Q se escribe:

$$Q(x) = \sum \lambda_i y_i^2, \quad r \text{ de los } \lambda_i, \text{ exactamente, son } \neq 0.$$

Podemos suponer

$$\lambda_1 \neq 0, \quad \lambda_2 \neq 0, \dots, \lambda_r \neq 0 \quad \text{y} \quad \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0.$$

Existen elementos $\alpha_i \in \mathbf{C}$ ($1 \leq i \leq r$) tales que $\alpha_i^2 = \lambda_i$. Hacemos:

$$e_i = \frac{f_i}{\alpha_i} \quad \text{para } 1 \leq i \leq r, \quad e_i = f_i \quad \text{para } i > r,$$

y designamos por x_1, \dots, x_n a las coordenadas de x en la base (e_1, \dots, e_n) . Para $1 \leq i \leq r$, se tiene $x_i = \alpha_i y_i$ de modo que:

$$Q(x) = x_1^2 + x_2^2 + \cdots + x_r^2. \quad \text{c.q.d.}$$

Incidentalmente, hemos demostrado:

XII.4.2 Para toda forma cuadrática no degenerada sobre \mathbf{C}^n , existe una base ortonormal.

Nota. En general, XII.4.1 y XII.4.2 permanecen verdaderos reemplazando \mathbf{C} por K , en donde K designa un cuerpo algebraicamente cerrado.

TEOREMA XII.4.3

Sea Q una forma cuadrática de rango r sobre \mathbf{R}^n . Entonces Q es equivalente a una forma del tipo:

$$x_1^2 + x_2^2 + \cdots + x_p^2 - x_{p+1}^2 - x_{p+2}^2 - \cdots - x_r^2,$$

en donde p es un entero que sólo depende de Q .

Este teorema se conoce con el nombre de la *ley de inercia de Sylvester*. Al par (p, q) , en donde p designa el número de cuadrados precedidos del signo $+$, y $q = r - p$ el número de cuadrados precedidos del signo $-$ (r designa el rango de Q), se le llama *signatura*, o *tipo* de la forma cuadrática Q .

(En consecuencia: para que dos formas cuadráticas sobre un \mathbf{R} -espacio vectorial de dimensión finita sean *equivalentes*, es necesario y suficiente que tengan la *misma signatura*.)

Demostración

a) Elijamos una base ortogonal para Q , por ejemplo (f_1, \dots, f_n) , en la cual

$$Q(\sum y_i f_i) = \sum_{i=1}^r a_i y_i^2 \quad (a_i \neq 0 \text{ para } 1 \leq i \leq r).$$

Podemos suponer que

$$a_i > 0 \quad \text{para } 1 \leq i \leq p \quad \text{y} \quad a_i < 0 \quad \text{para } p+1 \leq i \leq r.$$

Hacemos

$$\alpha_i = \sqrt{a_i} \quad \text{para } 1 \leq i \leq p, \quad \alpha_i = \sqrt{-a_i} \quad \text{para } p+1 \leq i \leq r,$$

$$e_i = \frac{f_i}{\alpha_i} \quad \text{para } 1 \leq i \leq r, \quad \text{y} \quad e_i = f_i \quad \text{para } i > r.$$

Entonces se tiene:

$$(1) \quad Q(\sum x_i e_i) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2.$$

b) Debemos demostrar que el entero p no depende de la base (e_i) en la que Q se escribe en la forma (1).

Para ello tomamos dos bases (e_1, \dots, e_n) , (e'_1, \dots, e'_n) , tales que

$$Q(\sum x_i e_i) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2,$$

$$Q(\sum x'_i e'_i) = x_1'^2 + \dots + x_{p'}^2 - x_{p'+1}'^2 - \dots - x_r'^2.$$

Sea F el subespacio de E engendrado por (e_1, \dots, e_p) .

G el subespacio de E engendrado por (e_{p+1}, \dots, e_n)

y, análogamente, definamos a partir de $e'_1, \dots, e'_{p'}, \dots, e'_n$, los subespacios F' y G' .

Para $x \in F \setminus \{0\}$ se tiene $Q(x) > 0$, y para $x \in G$, $Q(x) \leq 0$.

Análogamente, $Q(x) > 0$ para $x \in F' \setminus \{0\}$ y $Q(x) \leq 0$ para $x \in G'$.

Se deduce que $F' \cap G = \{0\}$, de donde $\dim(F') + \dim(G) \leq n$, o sea: $p' \leq p$. Análogamente se tiene: $p \leq p'$, de donde $p' = p$. c.q.d.

Observemos que el teorema XII.4.1 no es válido en el caso de \mathbf{R}^n . En este caso, las únicas formas cuadráticas Q que admiten una base ortonormal son las formas del tipo $(n, 0)$, caracterizadas por el hecho de que

$$Q(x) > 0 \quad \text{para } x \neq 0.$$

DEFINICIÓN XII.4.2

Una forma cuadrática sobre \mathbf{R}^n es **definida positiva** si es del tipo $(n, 0)$, **definida negativa** si es del tipo $(0, n)$. En general, a una forma cuadrática Q de rango r se le llama **positiva** si es del tipo $(r, 0)$, **negativa** si es del tipo $(0, r)$.

Una forma positiva (resp. negativa) se caracteriza por el hecho de que $Q(x) \geq 0$ (resp. $Q(x) \leq 0$) para todo x .

Estas nociones se extienden al caso de un \mathbf{R} -espacio vectorial cualquiera E : a una forma cuadrática Q sobre E se le llama *positiva* (resp. *definida positiva*) si se verifica $Q(x) \geq 0$ (resp. $Q(x) > 0$) para todo $x \in E \setminus \{0\}$.

Procedimiento práctico de reducción a una suma de cuadrados

Para hallar el rango de una forma cuadrática sobre \mathbf{R}^n o \mathbf{C}^n , se puede examinar su matriz respecto de una base. Pero este método no permite, en general, obtener la signatura de una forma cuadrática sobre \mathbf{R}^n . El *método de Gauss*, que expon-dremos ahora, permite obtener esta determinación. Es un procedimiento por recu-rrencia que proporciona directamente una descomposición, y permite escribir cual-quier forma cuadrática en la forma $\sum \lambda_i u_i^2$, en donde los u_i designan formas lineales

independientes con coeficientes reales. Entonces queda de manifiesto la signatura (p, q) de esta forma cuadrática, en la que p es el número de coeficientes $\lambda_i > 0$, y q , el número de coeficientes $\lambda_i < 0$.

Método de Gauss

Razonemos por recurrencia, y supongamos que toda forma cuadrática Q sobre K^{n-1} ($K = \mathbf{R}$ o \mathbf{C}), la sabemos expresar en la forma $Q = \sum \lambda_i u_i^2$, en donde los u_i son formas lineales independientes respecto de x_1, \dots, x_{n-1} .

Sea entonces $Q(x_1, \dots, x_n)$ una forma cuadrática sobre K^n .

Primer caso. Q contiene un término en x_1^2 y se puede escribir:

$$Q(x_1, \dots, x_n) = \lambda_1 x_1^2 + 2 R(x_2, \dots, x_n) x_1 + S(x_2, \dots, x_n),$$

con $\lambda_1 \neq 0$, $R(x_2, \dots, x_n)$ designa una forma lineal respecto de x_2, \dots, x_n , y $S(x_2, \dots, x_n)$ una forma cuadrática respecto de x_2, \dots, x_n .

Se escribe:

$$Q(x_1, \dots, x_n) = \lambda_1 \left(x_1 + \frac{R}{\lambda_1} \right)^2 - \frac{R^2}{\lambda_1} + S(x_2, \dots, x_n),$$

en donde R substituye a $R(x_2, \dots, x_n)$. Por la hipótesis de recurrencia, se tiene:

$$S(x_2, \dots, x_n) - \frac{R^2}{\lambda_1} = \sum_{i=2}^r \lambda_i u_i^2,$$

en donde los u_i son formas lineales independientes respecto de x_2, \dots, x_n . Si hacemos $u_1 = x_1 + \frac{R}{\lambda_1}$, la relación anterior se transforma en:

$$Q(x_1, \dots, x_n) = \sum_{i=1}^r \lambda_i u_i^2,$$

y los u_i son también formas independientes puesto que únicamente u_1 contiene a x_1 .

Con mayor precisión, la matriz de u_1, \dots, u_n en la base dual (e_i^*) de la base

canónica de K^n es de la forma $B = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \times & & & \\ \vdots & A & & \\ \times & & & \end{bmatrix}$, en donde A es la matriz de

u_2, \dots, u_n en (e_2^*, \dots, e_n^*) y se tiene $\det(B) = \det(A) \neq 0$.

Segundo caso. Q no contiene términos en x_1^2 . Entonces (numerando convenientemente los x_i) tendremos:

$$Q(x_1, \dots, x_n) = ax_1 x_2 + Rx_1 + Sx_2 + T(x_3, \dots, x_n), \quad a \neq 0,$$

en donde R y S son formas lineales en x_3, \dots, x_n , y T es una forma cuadrática en x_3, \dots, x_n .

Se escribe:

$$Q(x_1, \dots, x_n) = a \left[\left(x_1 + \frac{S}{a} \right) \left(x_2 + \frac{R}{a} \right) - \frac{RS}{a^2} \right] + T,$$

luego

$$\left(x_1 + \frac{S}{a} \right) \left(x_2 + \frac{R}{a} \right) = \frac{1}{4} (u_1^2 - u_2^2),$$

con:

$$u_1 = x_1 + x_2 + \frac{R}{a} + \frac{S}{a}, \quad u_2 = x_1 - x_2 + \frac{S}{a} - \frac{R}{a};$$

finalmente

$$T - \frac{RS}{a} = \sum_{i=3}^r \lambda_i u_i^2,$$

en donde los u_i ($3 \leq i \leq r$) son formas lineales independientes respecto de x_3, \dots, x_n .

Es claro que $u_1, u_2, u_3, \dots, u_r$ son linealmente independientes (únicamente u_1 y u_2 contienen a x_1 y a x_2 , y u_1, u_2 son independientes). Con precisión, la matriz de u_1, \dots, u_n en la base dual (e_i^*) de la base canónica de K^n es de la forma

$$B = \begin{bmatrix} 1 & 1 & 0 \dots 0 \\ 1 & -1 & 0 \dots 0 \\ \times & \times & \\ \vdots & \vdots & A \\ \times & \times & \end{bmatrix},$$

en donde A es la matriz de u_3, \dots, u_n en (e_3^*, \dots, e_n^*) , y se tiene:

$$Q(x_1, \dots, x_n) = \frac{a}{4} (u_1^2 - u_2^2) + \sum_{i=3}^r \lambda_i u_i^2,$$

con

$$\det(B) = -2 \det(A) \neq 0.$$

Este procedimiento permite, pues, escribir, paso a paso, Q en la forma enunciada. Proporciona la signatura de las formas sobre \mathbf{R}^n .

Ejemplos

1) Hallar el tipo de la forma con tres variables (sobre \mathbf{R}):

$$Q = yz + zx + xy.$$

Se escribe:

$$Q = xy + zx + zy = (x+z)(y+z) - z^2 = \frac{1}{4}(x+y+2z)^2 - \frac{1}{4}(x-y)^2 - z^2$$

Por lo tanto, Q es una forma de rango 3, de tipo (1, 2).

2) Hallar el tipo de la forma con 4 variables (sobre \mathbf{R}):

$$Q = x^2 + 2y^2 + 3z^2 - 2zt + tx + 3xy - yt;$$

se escribe:

$$Q = 3z^2 - 2tz + Q_1 = 3\left(z - \frac{t}{3}\right)^2 - \frac{t^2}{3} + Q_1,$$

$$Q_1 - \frac{t^2}{3} = x^2 + 2y^2 - \frac{t^2}{3} + tx + 3xy - yt = \left[x + \frac{1}{2}(t + 3y)\right]^2 - \frac{1}{4}(t + 3y)^2 + 2y^2 - \frac{t^2}{3} - yt,$$

$$2y^2 - \frac{t^2}{3} - yt - \frac{1}{4}(t + 3y)^2 = -\frac{1}{4}y^2 - \frac{7}{12}t^2 - \frac{5}{2}yt = -\frac{1}{4}(y + 5t)^2 + \frac{17}{3}t^2;$$

finalmente,

$$Q = 3\left[z - \frac{t}{3}\right]^2 + \left[x + \frac{1}{2}(t + 3y)\right]^2 - \frac{1}{4}[y + 5t]^2 + \frac{17}{3}t^2.$$

Por lo tanto, Q es una forma de rango 4, de signatura (3, 1).

Nota. Tiene interés tratar ante todo, si las hay, las variables que se presentan con mayor simplicidad.

§ XII.5 ESPACIO EUCLÍDEO

DEFINICIÓN XII.5.1

Un espacio euclídeo se define, dando:

- 1) un espacio vectorial E de dimensión finita sobre \mathbf{R} ;
- 2) una forma cuadrática Q definida positiva sobre E . A la forma bilineal B asociada a Q se le llama producto escalar asociado a Q .

También podemos definir un espacio euclídeo dando el espacio vectorial E y una forma bilineal no degenerada B tal que $B(x, x) \geq 0$ para todo $x \in E$; en esta forma, la definición se puede extender a los espacios vectoriales de dimensión infinita sobre \mathbf{R} y conduce a la noción de espacio prehilbertiano real (ver Cap. XIII, § 3). Algunas de las demostraciones que siguen permanecen válidas para estos nuevos espacios.

Ejemplos

1) En el espacio $E = \mathbf{R}^n$, la forma $(x, y) \mapsto \sum x_i y_i$ es un producto escalar, cuya forma cuadrática asociada es $x \mapsto \sum x_i^2$. Provisto de este producto escalar, E es entonces un espacio euclídeo designado por E_n . Todos los espacios euclídeos de dimensión n son isomorfos con él, puesto que una forma cuadrática definida positiva sobre un espacio de dimensión finita admite una base ortonormal (cf. § 4). El producto escalar de E_n se designa, en general, por $x \cdot y$.

Si x, y se representan por las matrices columna \mathcal{X}, \mathcal{Y} , el producto escalar $x \cdot y$ se identifica con el producto de matrices ${}^t\mathcal{X} \cdot \mathcal{Y}$ (este producto es una $(1, 1)$ -matriz, es decir, un escalar).

2) Sea $\mathcal{C}_{[0,1]}$ el \mathbf{R} -espacio vectorial de las aplicaciones *continuas*

$$f : [0, 1] \rightarrow \mathbf{R}.$$

La aplicación $(f, g) \mapsto \int_0^1 f(t) g(t) dt$, $\mathcal{C}_{[0,1]} \times \mathcal{C}_{[0,1]}$ en \mathbf{R} , es una forma bilineal definida positiva, puesto que,

$$(f \neq 0) \Rightarrow (f^2 \geq 0 \text{ y } f \neq 0) \Rightarrow \int_0^1 f^2(t) dt > 0.$$

Se trata, por lo tanto, un producto escalar que convierte a $\mathcal{C}_{[0,1]}$ en un espacio prehilbertiano real. Este ejemplo es fundamental en Análisis, en donde se consideran frecuentemente subespacios de $\mathcal{C}_{[0,1]}$ (por ejemplo, el subespacio de las funciones polinomios, o el subespacio engendrado por las funciones $(\sin nx)_{n \in \mathbf{N}}$, etc.).

Dual de un espacio euclídeo

Sea E un espacio euclídeo definido por la forma bilineal B . Según XII.3.2 toda forma lineal φ sobre E se escribe, de una manera y una sola, en la forma $\varphi(x) = B(x, y)$, en donde $y \in E$. Luego E^* es isomorfo a E .

Métrica asociada a un espacio euclídeo de dimensión finita n **TEOREMA XII.5.1**

Sea E un espacio euclídeo, Q la forma definida positiva de E , B la forma polar de Q . Si hacemos $N(x) = (Q(x))^{1/2}$ para $x \in E$, la función N es una **norma** sobre el espacio euclídeo E , es decir, verifica:

$$(1) \quad N(\lambda x) = |\lambda| N(x),$$

$$(2) \quad N(x) = 0 \Leftrightarrow x = 0,$$

$$(3) \quad N(x + y) \leq N(x) + N(y) \text{ para todo } x \in E, \text{ todo } y \in E \text{ y todo } \lambda \in \mathbf{R} \text{ (desigualdad triangular).}$$

A N se le llama la *norma euclídea* asociada a (E, Q) . Recordemos (cf. curso de Análisis, tomo 2 de la presente obra) que entonces se define sobre E una *estructura de espacio métrico* estableciendo:

$$d(x, y) = N(x - y).$$

La demostración de XII.5.1 se basa en la *desigualdad de Cauchy-Schwarz*, que vamos a establecer en primer lugar:

TEOREMA XII.5.2 (Desigualdad de Cauchy-Schwarz)

Con las notaciones que preceden, se tiene:

$$(4) \quad [B(x, y)]^2 \leq Q(x) Q(y),$$

verificándose la igualdad solamente cuando x, y son colineales.

Demostración de XII.5.2. Si $Q(x)$ y $Q(y)$ no son ambos nulos, por ejemplo, si $Q(y) \neq 0$, se tiene, para todo $\lambda \in \mathbf{R}$:

$$Q(x + \lambda y) \geq 0,$$

es decir, utilizando la relación $Q(u) = B(u, u)$:

$$(5) \quad Q(x) + 2\lambda B(x, y) + \lambda^2 Q(y) \geq 0.$$

El primer miembro de (5) es un trinomio en λ que sólo toma valores no negativos, luego su discriminante es ≤ 0 , es decir, que se verifica (4).

Si $Q(x) = Q(y) = 0$ se tiene $x = y = 0$ puesto que Q es definida positiva, y (4) también es verdadero.

Supongamos ahora que $[B(x, y)]^2 = Q(x) Q(y)$. El primer miembro de (5) tiene entonces una raíz doble en λ , necesariamente real (y de signo opuesto al de $B(x, y)$). Existe, pues, $\xi \in \mathbf{R}$ tal que $x + \xi y = 0$, de ahí la última afirmación del teorema. \square

Nota. Si expresamos (4) en una base ortonormal (e_i) , y hacemos $x = \sum x_i e_i$, $y = \sum y_i e_i$, se obtiene:

$$(\sum x_i y_i)^2 \leq (\sum x_i^2) (\sum y_i^2).$$

Demostración del teorema XII.5.1. Las relaciones (1) y (2) son evidentes; bastará comprobar (3). Puesto que los dos miembros de (3) son ≥ 0 , (3) es equivalente a la desigualdad obtenida elevando al cuadrado, que es:

$$(6) \quad Q(x + y) \leq Q(x) + Q(y) + 2(Q(x) Q(y))^{1/2}.$$

Puesto que $Q(x + y) = Q(x) + Q(y) + 2 B(x, y)$, (6) equivale a:

$$B(x, y) \leq (Q(x) Q(y))^{1/2},$$

lo cual se sigue de la desigualdad de Cauchy-Schwarz.

Nota. Si se tiene la igualdad: $N(x + y) = N(x) + N(y)$, se deduce:

$$B(x, y) = (Q(x) Q(y))^{1/2},$$

por lo tanto, en virtud de XII.5.2, existe un número λ (necesariamente positivo) que verifica $y = \lambda x$ o $x = \lambda y$. c.q.d.

Estructura euclídea inducida

Sea (E, Q) un espacio euclídeo cualquiera. Si F es un subespacio de E , la restricción de Q a F es una forma cuadrática definida positiva, que define sobre F una estructura de espacio euclídeo. Para esta estructura, el producto escalar es la restricción a $F \times F$ del producto escalar de E . A esta estructura se le llama *estructura euclídea inducida por E en F* . Cuando hablemos del *subespacio euclídeo F* , nos referiremos siempre al espacio euclídeo así definido.

Producto de dos espacios euclídeos

Sean (E_1, Q_1) , (E_2, Q_2) dos espacios euclídeos. En el espacio vectorial producto $E = E_1 \times E_2$ definimos la función Q :

$$x = (x_1, x_2) \mapsto Q(x) = Q_1(x_1) + Q_2(x_2).$$

Q es una forma cuadrática definida positiva sobre E , que convierte a E en un espacio euclídeo. Al espacio (E, Q) se le llama *espacio euclídeo producto* (o *suma directa*) de E_1 y E_2 , y se designa por $E_1 \oplus E_2$.

Si B_1 , B_2 y B designan los productos escalares asociados a Q_1 , Q_2 y Q , se tiene:

$$B(x, y) = B_1(x_1, y_1) + B_2(x_2, y_2) \quad \text{para } x = (x_1, x_2) \text{ e } y = (y_1, y_2).$$

De la misma manera se define el producto de m espacios euclídeos.

Subespacios de un espacio euclídeo de dimensión finita

Vamos a estudiar la familia de subespacios de un espacio euclídeo (E_n, Q) de dimensión n . Si F es uno de los subespacios, sabemos ya que

$$\dim(F) + \dim(F^\perp) = n \quad (\text{T. XII.3.3}).$$

Por otro lado, se tiene también $F \cap F^\perp = \{0\}$, pues todo elemento $x \in F \cap F^\perp$ es tal que $Q(x) = 0$, de donde $x = 0$. (Una forma definida positiva no posee ningún vector isótropo no nulo.) Hemos demostrado, pues, el teorema fundamental siguiente:

TEOREMA XII.5.3

|| Para todo espacio euclídeo (E_n, Q) de dimensión finita n , y para todo subespacio F de E_n , E_n es suma directa de F y del ortogonal F^\perp a F .

A F^\perp se le llama **suplementario ortogonal** de F .

Los papeles de F y F^\perp son simétricos, puesto que $(F^\perp)^\perp = F$ (T. XII.3.3). Sea ahora x un elemento de E_n , que descomponemos sobre los espacios F y F^\perp :

$$x = y + z \quad y \in F, \quad z \in F^\perp.$$

Calculamos $Q(x)$:

$$Q(x) = Q(y + z) = Q(y) + Q(z) + 2B(y, z) = Q(y) + Q(z),$$

puesto que $B(y, z) = 0$.

La relación:

$$(7) \quad Q(x) = Q(y) + Q(z)$$

es la forma general del *teorema de Pitágoras*. Se puede expresar también diciendo que el espacio euclídeo E_n es isomorfo al producto de los subespacios euclídeos F

y F^\perp . Podemos generalizar este resultado: decimos que dos subespacios F y G son *ortogonales* si $F \subset G^\perp$, lo cual significa que todo vector de F es ortogonal a todo vector de G . La relación: $F \subset G^\perp$ es, entonces, equivalente a la relación: $G \subset F^\perp$.

TEOREMA XII.5.4

Sean (E_n, Q) un espacio euclídeo de dimensión n , y F_1, \dots, F_k subespacios de E_n **ortogonales dos a dos**, tales que el espacio vectorial E_n sea la suma directa de los F_i .

Si $x = \sum_{i=1}^k x_i$ es la descomposición de un elemento x de E_n sobre los F_i , se tiene:

$$Q(x) = \sum_{i=1}^k Q(x_i).$$

Demostración. Por recurrencia sobre k , con la ayuda de (7). c.q.d.

Luego el espacio euclídeo E_n es la suma directa de los subespacios euclídeos F_i .

Caso particular

Sea (e_1, \dots, e_n) una base ortonormal del espacio euclídeo (E_n, Q) , y sea F_i el subespacio engendrado por e_i . E_n es la suma directa de los subespacios euclídeos F_i . luego, si $x = \sum x_i e_i$, se tiene:

$$Q(x) = \sum_{i=1}^n Q(x_i e_i) = \sum x_i^2 Q(e_i) = \sum x_i^2, \quad \text{por una parte.}$$

Por otra parte:

$$B(x, e_i) = \sum_{j=1}^n B(x_j e_j, e_i) = x_i Q(e_i) = x_i,$$

puesto que $B(e_j, e_i) = \delta_{ij}$. Podemos escribir entonces la siguiente relación importante (llamada de *Parseval*):

$$(8) \quad \boxed{Q(x) = \sum_{i=1}^n [B(x, e_i)]^2},$$

válida para toda base ortonormal de E_n .

● En adelante, fijada la forma Q , el producto escalar $B(x, y)$ del espacio euclídeo (E, Q) se designará simplemente por $x \cdot y$, y se escribirá $|x| = (x \cdot x)^{1/2}$.

DEFINICIÓN XII.5.2

Se dice que los vectores (e_1, \dots, e_p) del espacio euclídeo E_n forman un **sistema ortonormal** si son dos a dos ortogonales y si su norma es igual a 1, en otras palabras, si $e_i \cdot e_j = \delta_{ij}$ (símbolo de Kronecker) cualesquiera que sean $i, j = 1, 2, \dots, p$.

Propiedades

Un sistema ortonormal es libre. En efecto, una relación de la forma:

$$\sum_{i=1}^p \lambda_i e_i = 0 \quad \text{implica} \quad \left(\sum_{i=1}^p \lambda_i e_i \right) \cdot e_j = \sum_{i=1}^p \lambda_i (e_i \cdot e_j) \\ = \sum_{i=1}^p \lambda_i \delta_{ij} = \lambda_j = 0 \quad \text{para } 1 \leq j \leq p.$$

En particular, se tiene necesariamente $p \leq n$. Cuando $p = n = \dim(E_n)$, el sistema (e_1, \dots, e_n) es ortonormal si, y sólo si, es una base ortonormal de E_n .

Sean F, G dos subespacios de E_n , (e_1, \dots, e_p) un sistema ortonormal de F , (f_1, \dots, f_q) un sistema ortonormal de G . Si F y G son ortogonales (e.d. $G \subset F^\perp$, lo que equivale a $F \subset G^\perp$), el sistema

$$(e_1, \dots, e_p ; f_1, \dots, f_q)$$

es también ortonormal.

Toda parte de un sistema ortonormal es un sistema ortonormal. En general, sea (e_1, \dots, e_p) un sistema ortonormal, que engendre el subespacio F , y sea (J_1, J_2, \dots, J_k) una *partición* de \mathbf{N}_p^* . Si ponemos

$$F_i = \text{Vect} [(e_j)_{j \in J_i}],$$

los F_i son ortogonales dos a dos, y F es la suma directa de los F_i . (El lector observará la analogía de estas propiedades con las de las partes libres de un espacio vectorial.)

Sabemos que todo espacio euclídeo E_n de dimensión finita n , admite una base ortonormal. Por consiguiente, si F es un subespacio vectorial de E_n , F admite una base ortonormal (e_1, \dots, e_p) (en donde $p = \dim(F)$). Asimismo, F^\perp admite una base ortonormal (e_{p+1}, \dots, e_n) . Según las observaciones anteriores, $(e_1, \dots, e_p, \dots, e_n)$ es una base ortonormal de E_n . Podemos enunciar, pues:

XII.5.5 *Todo sistema ortonormal de un espacio euclídeo de dimensión finita se puede completar de forma que se obtenga una base ortonormal de todo el espacio.*

En la práctica, se utiliza el procedimiento que sigue:

XII.5.6 (Procedimiento de ortonormalización de Schmidt.) Sea (v_1, \dots, v_n) una base cualquiera del espacio euclídeo (E_n, Q) . Existe una base ortonormal (e_1, \dots, e_n) única de E_n , que verifica las siguientes condiciones:

- a) Para todo entero p , $e_p \cdot v_p > 0$.
- b) Para todo entero p , $\text{Vect}(e_1, \dots, e_p) = \text{Vect}(v_1, \dots, v_p)$.

Demostración. Construiremos, por recurrencia sobre p , los vectores e_1, \dots, e_n de forma que (e_1, \dots, e_p) constituya un sistema ortonormal para todo p , que verifique $e_k \cdot v_k > 0$ para $k \leq p$, y que $\text{Vect}(e_1, \dots, e_p) = \text{Vect}(v_1, \dots, v_p)$; y veremos que los e_i están unívocamente determinados.

Necesariamente $e_1 = \frac{v_1}{|v_1|}$ (haciendo, según hemos convenido, $|v| = (Q(v))^{1/2}$).

Supongamos que ya hemos construido (e_1, \dots, e_p) : e_{p+1} debe ser de la forma:

$$e_{p+1} = \mu v_{p+1} + \lambda_1 e_1 + \dots + \lambda_p e_p.$$

Las condiciones $e_{p+1} \cdot e_i = 0$ ($1 \leq i \leq p$) nos dan: $\lambda_i = -\mu v_{p+1} \cdot e_i$.

Por lo tanto, se tiene $e_{p+1} = \mu w_{p+1}$, con

$$(9) \quad w_{p+1} = v_{p+1} - \sum_{i=1}^p (v_{p+1} \cdot e_i) e_i \quad \text{y} \quad \mu > 0$$

(ya que $1 = |e_{p+1}|^2 = \mu w_{p+1} \cdot e_{p+1} = \mu v_{p+1} \cdot e_{p+1}$ y $v_{p+1} \cdot e_{p+1} > 0$).

De donde, necesariamente, $\mu = \frac{1}{|w_{p+1}|}$.

Recíprocamente el vector w_{p+1} definido por (9) es no nulo, pues v_{p+1} no pertenece a $\text{Vect}(v_1, \dots, v_p) = \text{Vect}(e_1, \dots, e_p)$. Entonces se comprueba sin ninguna dificultad que el vector $e_{p+1} = \frac{w_{p+1}}{|w_{p+1}|}$ verifica las condiciones requeridas. c.q.d.

Nota. XII.5.6 se puede expresar matricialmente como sigue: toda matriz invertible M se puede poner de manera única en la forma $M = AT$, en donde A es ortogonal (ver § 7) y en donde T es triangular superior con coeficientes diagonales > 0 .

Ejemplos

1) En \mathbf{R}^3 euclídeo, ortonormalizar la base

$$v_1(1, 1, 1), \quad v_2(1, 1, 0), \quad v_3(1, 0, 0).$$

Siguiendo el método de XII.5.7 se obtienen sucesivamente:

$$e_1 = \frac{v_1}{\sqrt{3}}, \quad w_2 = v_2 - (v_2 \cdot e_1) e_1, \quad w_2 = \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3} \right), \quad e_2 = \frac{\sqrt{3}}{\sqrt{2}} w_2,$$

$$w_3 = v_3 - (v_3 \cdot e_1) e_1 - (v_3 \cdot e_2) e_2, \quad e_3 = \frac{w_3}{|w_3|},$$

de donde resulta que la base buscada es:

$$e_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right), \quad e_2 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}} \right), \quad e_3 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0 \right)$$

2) Polinomios de Legendre.

Son los polinomios: $L_n = \frac{d^n}{dX^n} ((X^2 - 1)^n)$. Consideremos el espacio vectorial de las funciones polinomio en $[-1, 1]$, con el producto escalar B definido por

$$B(P, Q) = \int_{-1}^1 P(x) Q(x) dx.$$

Se obtiene un espacio prehilbertiano real. Vamos a ver que los polinomios L_n son ortogonales dos a dos, y que para todo

$$n \geq 0, \quad \text{Vect}(L_0, \dots, L_n) = \text{Vect}(1, X, \dots, X^n).$$

La segunda de dichas afirmaciones es prácticamente evidente. Como L_n es exactamente de grado n , la matriz de (L_0, \dots, L_n) en la base $(1, X, \dots, X^n)$ es triangular superior, con elementos $\neq 0$ en la diagonal. Para demostrar la primera, es suficiente observar que, por medio de una integración por partes, se obtiene (con $p \leq n - 1$):

$$\begin{aligned} I_p = \int_{-1}^1 \left(\frac{d^n}{dx^n} [(x^2 - 1)^n] \right) x^p dx &= \left[x^p \frac{d^{n-1}}{dx^{n-1}} [(x^2 - 1)^n] \right]_{-1}^1 \\ &\quad - \int_{-1}^1 p x^{p-1} \frac{d^{n-1}}{dx^{n-1}} [(x^2 - 1)^n] dx. \end{aligned}$$

Puesto que 1 y -1 son raíces de orden n de $(x^2 - 1)^n$, son raíces de orden $n - p$ de $\frac{d^p}{dx^p} ((x^2 - 1)^n)$ para $0 \leq p \leq n - 1$. En la fórmula anterior, la parte ya integrada desaparece, de donde:

$$I_p = - \int_{-1}^1 p x^{p-1} \frac{d^{n-1}}{dx^{n-1}} [(x^2 - 1)^n] dx,$$

y, por recurrencia, para todo $k \leq p$,

$$I_p = (-1)^k \int_{-1}^1 p(p-1) \dots (p-k+1) x^{p-k} \frac{d^{n-k}}{dx^{n-k}} [(x^2-1)^n] dx.$$

En particular, para $k = p$,

$$I_p = (-1)^p p! \int_{-1}^1 \frac{d^{n-p}}{dx^{n-p}} [(x^2-1)^n] dx = (-1)^p p! \left[\frac{d^{n-p-1}}{dx^{n-p-1}} [(x^2-1)^n] \right]_{-1}^1 = 0.$$

Puesto que $I_p = 0$, se tiene, por linealidad: $\int_{-1}^1 P(x) \cdot L_n(x) dx = 0$ para todo polinomio P de grado $\leq n-1$. De donde

$$\int_{-1}^1 L_p(x) L_n(x) dx = 0 \quad \text{para } p \leq n-1,$$

según habíamos enunciado.

Extensión de la desigualdad de Cauchy-Schwarz

Sean E un \mathbf{R} -espacio vectorial y Q una forma cuadrática sobre E , de forma polar B . Se dice que Q es *positiva* (o que B es positiva) si $\forall x \in E$ $Q(x) \geq 0$, *definida positiva* si es positiva y si $(Q(x) = 0) \Rightarrow (x = 0)$. Entonces se tiene:

XII.5.7 Para toda forma Q positiva sobre E , se tiene

$$\| \quad \forall x \in E, \forall y \in E \quad [B(x, y)]^2 \leq Q(x) Q(y).$$

Demostración. La función $\mathbf{R} \rightarrow \mathbf{R}$, $\lambda \mapsto Q(x + \lambda y)$, toma sus valores en \mathbf{R}_+ . Luego

$$Q(x + \lambda y) = Q(x) + 2\lambda B(x, y) + \lambda^2 Q(y).$$

Si $Q(y) > 0$, se tiene un trinomio que sólo toma valores positivos, de donde

$$[B(x, y)]^2 - Q(x) Q(y) \leq 0.$$

Si $Q(y) = 0$, la función afín $\lambda \mapsto Q(x) + 2\lambda B(x, y)$ debe tener signo constante, por lo tanto $B(x, y) = 0$. c.q.d.

COROLARIO

$\|$ Sea Q una forma cuadrática *positiva* sobre E . Para que Q sea *no degenerada*, es necesario y suficiente que sea *definida positiva*.

Demostración. Sea $x \in E \setminus \{0\}$; para toda $y \in E$, se tiene:

$$[B(x, y)]^2 \leq Q(x) Q(y), \text{ luego } (Q(x) = 0) \Rightarrow (\forall y \in E, B(x, y) = 0).$$

Por lo tanto, si Q no es definida positiva, Q es degenerada.

Recíprocamente, si Q es definida positiva, y si $x \in E \setminus \{0\}$, este elemento no puede ser ortogonal a E , ya que $Q(x) > 0$. Por lo tanto Q es no degenerada. c.q.d.

§ XII.6 PROYECCIONES Y SIMETRÍAS

En lo que sigue consideraremos un espacio euclídeo E_n de dimensión finita n . El producto escalar de dos elementos x, y de E_n lo designaremos por $x \cdot y$, su distancia euclídea por $d(x, y)$. La norma de $x \in E_n$ se designará por $|x|$. Entonces se tiene:

$$d(x, y) = |x - y|.$$

Recordemos (cf. Curso de Análisis) que la distancia de un punto $a \in E_n$ a una parte A de E_n es, por definición, el número positivo o nulo $d(a, A)$ definido por

$$d(a, A) = \inf_{x \in A} d(a, x).$$

Por otro lado, diremos que dos subespacios afines H, K de E_n son ortogonales si sus direcciones H_0, K_0 son ortogonales (recordemos que la dirección H_0 de un subespacio afín H de E_n es el subespacio vectorial de E_n respecto del cual H es un trasladado).

TEOREMA XII.6.1

Sea E_n un espacio euclídeo de dimensión n , y H un subespacio afín de E_n . Para cada punto a de E existe un punto $q(a)$ y sólo uno, tal que $q(a) \in H$ y $d(a, H) = d[a, q(a)]$. A este punto $q(a)$ se le llama **proyección ortogonal** de a sobre H , pues es el único punto b de H tal que $b - a$ es ortogonal a H . Finalmente, la aplicación $a \mapsto q(a)$ de E_n en H es una aplicación afín y epiyectiva, llamada **proyección ortogonal** sobre H .

Demostración. Designemos por h un punto cualquiera de H , tal que H sea el conjunto $h + H_0$. Por otra parte, sea (e_1, \dots, e_r) una base ortonormal de H_0 . Busquemos primeramente un punto $x \in H$ tal que: $x - a \in H_0^\perp$.

Si hacemos: $x = h + \sum_{i=1}^p \lambda_i e_i$, la condición $x - a \in H_0^\perp$ equivale a

$$\left(h - a + \sum_{i=1}^p \lambda_i e_i \right) \cdot e_j = 0 \quad \text{para } 1 \leq j \leq p,$$

por lo tanto a $e_j \cdot (h - a) + \lambda_j = 0$, $\lambda_j = (a - h) \cdot e_j$. Por lo tanto, el único punto que nos conviene es:

$$x = q(a) = h + \sum_{j=1}^p [(a - h) \cdot e_j] e_j.$$

Calculemos ahora la distancia $d(a, x)$ cuando $x \in H$ y $x \neq q(a)$. Se tiene

$$x = q(a) + x_0, \quad \text{con } x_0 \in H_0 \quad \text{y} \quad x_0 \neq 0.$$

$$[d(a, x)]^2 = |a - x|^2 = |a - q(a) - x_0|^2 = |c - x_0|^2, \quad \text{con } c = a - q(a).$$

Pero sabemos que $x_0 \in H_0$, y que $c = a - q(a) \in H_0^\perp$. Según XII.5.4, tenemos entonces:

$$|c - x_0|^2 = |c|^2 + |x_0|^2 > |c|^2 \quad (\text{puesto que } x_0 \neq 0);$$

y $d(a, x)$ (en donde x recorre H) es mínimo cuando $x = q(a)$.

Finalmente, la fórmula:

$$q(a) = h + \sum_{j=1}^p [(a - h) \cdot e_j] e_j$$

prueba que $a \mapsto q(a)$ es una aplicación afín. Si $a \in H$, $q(a) = a$, por lo tanto q es epiyectiva. c.q.d.

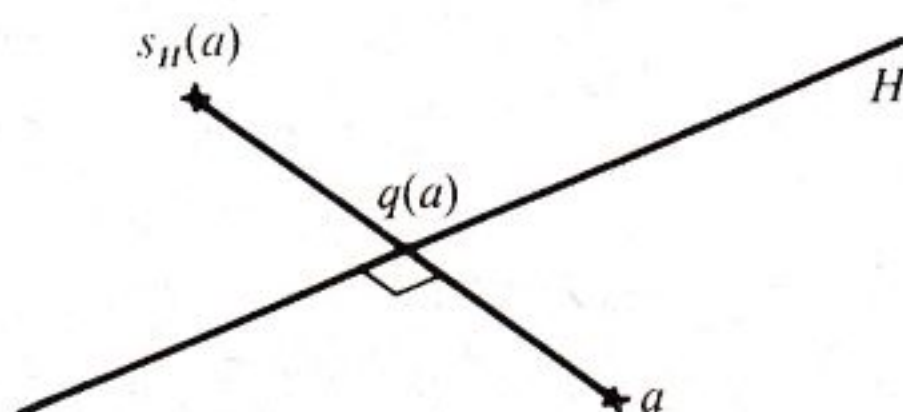
Nota. La imagen recíproca de $b \in H$ por q es el subespacio afín $b + L_0$, en donde L_0 designa el suplementario ortogonal H_0^\perp de H_0 .

Simetría respecto de un subespacio afín H de E_n

DEFINICIÓN XII.6.1

Si H designa un subespacio afín de E_n , y q la proyección ortogonal sobre H , la **simetría respecto de H** es la aplicación

$$a \mapsto s_H(a) = 2q(a) - a.$$



Es, por lo tanto, una aplicación afín.

Para todo $a \in E_n$, $s_H(a)$ es el único punto de E_n que verifica las siguientes condiciones: $s_H(a) - a$ es ortogonal a H_0 y el punto medio de $[a, s_H(a)]$ pertenece a H . (H_0 designa la dirección de H .)

Más adelante veremos que las simetrías respecto a un *hiperplano* desempeñan un papel importante en el estudio de las isometrías de E_n (cf. Cap. XIII).

Obsérvese que $s_h \circ s_h$ es la aplicación idéntica de E_n . Por lo que s_h es una *biyección* de E_n en sí mismo.

Hiperplano medio entre dos puntos

Sean a, b dos puntos *distintos* de E_n . Buscamos el conjunto de puntos $x \in E_n$ tales que $d(x, a) = d(x, b)$. Esta relación equivale a

$$|x - a|^2 = |x - b|^2, \text{ o sea } (x - a) \cdot (x - a) = (x - b) \cdot (x - b),$$

de donde, desarrollando:

$$(1) \quad 2(b - a) \cdot x = b^2 - a^2.$$

El conjunto de los $x \in E_n$ que verifican (1) constituye un hiperplano afín H , cuya dirección la forma el hiperplano vectorial H_0 de ecuación $(b - a) \cdot x = 0$. H_0 es precisamente el suplementario ortogonal de la recta engendrada por el vector $b - a$.

Según (1), H contiene el punto $\frac{a + b}{2}$, es decir, el punto medio de $[a, b]$.

DEFINICIÓN XII.6.2

Sean a, b dos puntos distintos del espacio euclídeo E de dimensión finita n .
Al hiperplano afín H de ecuación $(b - a) \cdot x = \frac{1}{2}(b^2 - a^2)$ (conjunto de puntos equidistantes de a y b) se le llama **hiperplano medio** del segmento $[a, b]$.

Observemos que entonces a y b son simétricos respecto de H .

§ XII.7 GRUPO ORTOGONAL, EL GRUPO ORTOGONAL REAL

De momento, el cuerpo base K es cualquiera. Sea Q una forma cuadrática cualquiera sobre el espacio vectorial E , y B su forma polar. Los automorfismos lineales φ de E tales que

$$(1) \quad \forall x \in E, \forall y \in E \quad B[\varphi(x), \varphi(y)] = B[x, y]$$

forman un grupo, pues si φ y ψ verifican (1), se tiene:

$$B[\varphi \circ \psi(x), \varphi \circ \psi(y)] = B[\psi(x), \psi(y)] = B[x, y],$$

luego $\varphi \circ \varphi$ verifica (1); y (si hacemos $x = \varphi(u)$, $y = \varphi(v)$)

$$B[\varphi^{-1}(x), \varphi^{-1}(y)] = B[u, v] = B[\varphi(u), \varphi(v)] = B[x, y],$$

luego φ^{-1} verifica (1).

Además, la identidad $\text{Id}(E)$ verifica (1). De ahí nuestra afirmación.

Según la fórmula: $B(x, y) = \frac{1}{2} [Q(x+y) - Q(x) - Q(y)]$, la relación (1) equivale a:

$$\forall x \in E \quad Q[\varphi(x)] = Q(x).$$

Con otras palabras, *es equivalente afirmar que el automorfismo φ conserva la forma bilineal B , o que conserva la forma cuadrática Q .*

DEFINICIÓN XII.7.1

Sea Q una forma cuadrática sobre un espacio vectorial E . Al subgrupo del grupo lineal $\text{GL}(E)$ formado por los automorfismos φ que conservan la forma Q , se le llama **grupo ortogonal de la forma Q** , y se designa por $O(Q)$. A sus elementos se le llama **ortogonales respecto de Q** .

Es claro que si sabemos clasificar las formas cuadráticas definidas sobre el espacio E , sabremos igualmente clasificar los grupos ortogonales sobre E .

Caso de un espacio E de dimensión finita n (cuerpo base K cualquiera, conmutativo y de característica $\neq 2$)

Sea Q una forma cuadrática *no degenerada* sobre E , de forma polar B . Entonces podemos establecer las siguientes precisiones acerca del grupo ortogonal $O(Q)$:

— Todo $\varphi \in \mathcal{L}_K(E)$ tal que $\forall x \in E \quad Q(\varphi(x)) = Q(x)$ es un automorfismo (luego $\varphi \in O(Q)$).

En efecto, un tal φ es entonces biyectivo (por lo tanto epiyectivo), pues se tiene:

$$\forall x \in E, \forall y \in E \quad B(\varphi(x), \varphi(y)) = B(x, y).$$

De donde: $(\varphi(x) = 0) \Rightarrow (\forall y \in E, B(x, y) = 0)$, y como Q es no degenerada, $\varphi(x) = 0$ implica pues $x = 0$.

— Un elemento de $O(Q)$ se puede caracterizar por su matriz en una base. Con precisión:

XII.7.1 Sean $\mathcal{B} = (e_1, \dots, e_n)$ una base de E y A la matriz de B en \mathcal{B} . Para que un endomorfismo $u \in \mathcal{L}_K(E)$ sea un elemento de $O(Q)$ es necesario y suficiente que su matriz M_u en \mathcal{B} verifique

$${}^t M_u A M_u = A.$$

En consecuencia, el conjunto de las matrices $M \in M_n(K)$ que verifican ${}^t M A M = A$ forma un subgrupo Γ_A de $GL_n(K)$, y la aplicación $u \mapsto M_u$, $O(Q) \rightarrow \Gamma_A$ es un isomorfismo de grupos.

Demostración. Si designamos por \mathcal{X} la matriz columna de las coordenadas en \mathcal{B} de un elemento $x \in E$ arbitrario, vemos que las relaciones:

$$\forall x : Q(\varphi(x)) = Q(x) \quad \text{y} \quad \forall \mathcal{X} : {}^t \mathcal{X} {}^t M_u A M_u \mathcal{X} = {}^t \mathcal{X} A \mathcal{X}$$

son equivalentes. Luego la segunda relación equivale a ${}^t M_u A M_u = A$, de donde se sigue la primera afirmación del teorema.

La segunda afirmación se deduce inmediatamente teniendo en cuenta el hecho de que $u \mapsto M_u$ es un isomorfismo de $GL_K(E)$ en $GL_n(K)$. c.q.d.

XII.7.1 es particularmente interesante cuando \mathcal{B} es **ortogonal** (e.d. A **diagonal**). Particularizando más, cuando la base \mathcal{B} es **ortonormal** (e.d. $A = I_n$) al grupo Γ_A se le llama **grupo de las matrices ortogonales de orden n con coeficientes en K** .

Ejemplo

Si \mathcal{B} es *ortogonal*, y si

$$M_u = \begin{bmatrix} \varepsilon_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \varepsilon_n \end{bmatrix} \quad \text{con} \quad \forall i \quad \varepsilon_i = \pm 1,$$

es evidente que $M_u \in \Gamma_A$, luego $u \in O(Q)$.

De la relación ${}^t M_u A M_u = A$, se deduce $[\det(M_u)]^2 \det(A) = \det(A)$, y puesto que $\det(A) \neq 0$, $[\det(M_u)]^2 = 1 = (\det(u))^2$.

De todo lo cual se deduce:

XII.7.2. La aplicación $d : u \mapsto \det(u)$ es un homomorfismo epiyectivo de $O(Q)$ en el subgrupo $G = \{-1, +1\}$ de K^* .

Demostración. Es inmediato que d está bien definida y que es un homomorfismo. Para terminar la demostración, es suficiente establecer la existencia de un $u \in O(Q)$ tal que $\det(u) = -1$. Sea, entonces, $\mathcal{B} = (e_1, \dots, e_n)$ una base *ortogonal* para Q .

El endomorfismo u de E tal que $u(e_i) = -e_i$ y $u(e_j) = e_j$ para $j \neq i$ satisface la condición. c.q.d.

Se da la siguiente

DEFINICIÓN XII.7.2

$\left\{ \begin{array}{l} \text{Se llama } \mathbf{grupo\ especial\ ortogonal\ de\ } Q, \text{ y se designa por } SO(Q) \\ \text{al subgrupo de } O(Q) \text{ formado por los } u \in O(Q) \text{ tales que } \det(u) = 1. \end{array} \right.$

Con otras palabras, $SO(Q)$ es el *núcleo* de d .

Según XII.7.2, el grupo cociente $O(Q)/SO(Q)$ (que está definido, puesto que $SO(Q)$ es un subgrupo normal de $O(Q)$), es isomorfo a G . Luego

$$[O(Q) : SO(Q)] = 2.$$

Dejamos al lector el trabajo de demostrar que *si dos formas cuadráticas son equivalentes sobre E , sus grupos ortogonales (resp. ortogonales especiales) son isomorfos.*

Grupo ortogonal real

En particular, consideremos una forma cuadrática Q *definida positiva* sobre un espacio vectorial real E de dimensión n . Sabemos que Q es equivalente a la forma $\sum_{i=1}^n x_i^2$. Por otra parte, si (e_i) es una base ortonormal de E , se tiene:

$$Q(\sum x_i e_i) = \sum x_i^2.$$

Luego el grupo $O(Q)$ es isomorfo al grupo ortogonal de la forma cuadrática $\sum x_i^2$ sobre el espacio \mathbf{R}^n .

DEFINICIÓN XII.7.3

$\left\{ \begin{array}{l} \text{Se llama } \mathbf{grupo\ ortogonal\ real\ de\ orden\ } n \text{ (o } \mathbf{grupo\ euclídeo\ de} \\ \mathbf{orden\ } n) \text{ y se designa por } O(n, \mathbf{R}), \text{ al grupo ortogonal de la forma} \\ \text{cuadrática } \sum_{i=1}^n x_i^2 \text{ sobre el espacio } \mathbf{R}^n. \end{array} \right.$

Con la ayuda de la base canónica (e_i) de \mathbf{R}^n , identificamos $GL(\mathbf{R}^n)$ con el grupo $GL(n, \mathbf{R})$ de las matrices cuadradas invertibles de orden n con elementos reales. Vamos a caracterizar los elementos $\varphi \in O(n, \mathbf{R})$ por medio de sus matrices M_φ .

Para ello designaremos por x, y a dos elementos de E y por \mathcal{X}, \mathcal{Y} a las matrices columna de sus coordenadas. La matriz de la forma $x.y = \sum_{i=1}^n x_i y_i$ en la base (e_i) es I_n , por lo que la fórmula de la fórmula (4) del § 1 resulta:

$$x.y = {}^t\mathcal{X} \cdot I_n \cdot \mathcal{Y} = {}^t\mathcal{X} \cdot \mathcal{Y}.$$

Por otra parte, si $x' = \varphi(x)$, e $y' = \varphi(y)$, se tiene $\mathcal{X}' = M_\varphi \mathcal{X}$, $\mathcal{Y}' = M_\varphi \mathcal{Y}$. De donde

$$x'.y' = {}^t\mathcal{X}' \cdot \mathcal{Y}' = {}^t\mathcal{X} \cdot ({}^tM_\varphi \cdot M_\varphi) \cdot \mathcal{Y}.$$

El endomorfismo φ conserva el producto escalar si, y sólo si, para todo $x \in E$ y todo $y \in E$, se tiene:

$${}^t\mathcal{X} I_n \mathcal{Y} = {}^t\mathcal{X} \cdot {}^tM_\varphi \cdot M_\varphi \cdot \mathcal{Y},$$

lo que equivale a ${}^tM_\varphi \cdot M_\varphi = I_n$ (pues estas dos matrices están asociadas a la misma forma bilineal). Inversamente, la relación ${}^tM_\varphi \cdot M_\varphi = I_n$ implica $\det(M_\varphi) \neq 0$, luego φ es invertible. Podemos enunciar, por lo tanto, el:

TEOREMA XII.7.3

El grupo ortogonal $O(n, \mathbf{R})$ es canónicamente isomorfo al grupo multiplicativo de las matrices cuadradas M de orden n sobre \mathbf{R} tales que:

$$(3) \quad {}^tM \cdot M = I_n.$$

(A estas matrices se les llama **ortogonales** con coeficientes reales.)

Desarrollamos (3) escribiendo $M = [a_{ij}]_{1 \leq i, j \leq n}$. Se obtiene:

$$(4) \quad \sum_{i=1}^n a_{ij} a_{ik} = \delta_{jk} \quad 1 \leq j, k \leq n, \quad \text{en donde } \delta_{jk} \text{ es el símbolo de Kronecker.}$$

Se ve inmediatamente que la relación ${}^tM \cdot M = I_n$ equivale a $M \cdot {}^tM = I_n$. Con otras palabras:

Si la matriz M es ortogonal, su traspuesta es ortogonal.

El sistema de las $\frac{n(n+1)}{2}$ relaciones (4) es, por lo tanto, *equivalente* al sistema

$$(4') \quad \sum_{i=1}^n a_{ji} a_{ki} = \delta_{jk} \quad (j, k = 1, 2, \dots, n).$$

Grupo $SO(n)$

Si tomamos determinantes en los dos miembros de (3), obtenemos

$$(5) \quad [\det(M)]^2 = 1, \text{ de donde } \det(M) = \pm 1.$$

La aplicación $M \mapsto \det(M)$ es un homomorfismo δ de $O(n, \mathbf{R})$ en el grupo multiplicativo $\{-1, +1\}$. Demostremos que este homomorfismo es epiyectivo, y para ello es suficiente observar que la matriz

$$M = \begin{bmatrix} 1 & 0 & \dots & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & \\ \vdots & & & \vdots \\ 0 & \dots & & -1 \end{bmatrix}$$

(que representa la simetría respecto del hiperplano $x_n = 0$) pertenece a $O(n, \mathbf{R})$ y tiene por determinante -1 .

Según la teoría de grupos, resulta que el núcleo de δ es un subgrupo de $O(n, \mathbf{R})$ de índice 2.

DEFINICIÓN XII.7.4

$\left\{ \begin{array}{l} \text{Al subgrupo de los } \varphi \in O(n, \mathbf{R}), \text{ cuyo determinante vale } +1, \text{ se le llama} \\ \text{grupo ortogonal especial de orden } n, \text{ y se designa por } SO(n, \mathbf{R}). \end{array} \right.$

Evidentemente, $SO(n, \mathbf{R})$ es un subgrupo normal de $O(n, \mathbf{R})$ (es el núcleo de un homomorfismo de grupos).

Como ejemplo, vamos a determinar los grupos $O(1)$, $SO(1)$, $O(2)$ y $SO(2)$. En el capítulo XIII estudiaremos más detalladamente los grupos $SO(n)$ y $O(n)$.

● $O(1)$ está formado por las matrices M de orden 1 tales que $M^t M = I_1$, es decir, los elementos $a \in \mathbf{R}$ tales que $a^2 = 1$. $O(1)$ es, por lo tanto, el grupo con dos elementos $\{-1, +1\}$. El elemento -1 representa la simetría respecto del origen, $+1$ la identidad, y $SO(1)$ se reduce a $\{1\}$.

● $O(2)$ está formado por las matrices $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tales que

$$(6) \quad \begin{aligned} a^2 + b^2 &= c^2 + d^2 = 1, \\ ac + bd &= 0, \quad ad - bc = \pm 1. \end{aligned}$$

Primer caso: $ad - bc = 1$. Las matrices correspondientes forman el grupo $SO(2)$. De $a^2 + b^2 = 1 = ad - bc$, se obtiene $a(a - d) + b(b + c) = 0$, de donde

se sigue la existencia de un $\lambda \in \mathbf{R}$ tal que $d = a - \lambda b$, $c = -\lambda a - b$ (ya que $a^2 + b^2 = 1$, a y b no son ambos nulos). Llevando estos valores de d y c a $ac + bd = 0$, se obtiene

$$-\lambda(a^2 + b^2) = 0, \text{ por tanto, } \lambda = 0, \text{ } d = a \text{ y } c = -b.$$

$\text{SO}(2)$ está formado, por lo tanto, por las matrices de la forma:

$$M = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \text{ en donde } a^2 + b^2 = 1$$

Existe un $\theta \in \mathbf{R}$ tal que $a = \cos \theta$, $b = \sin \theta$ (cf. Curso de Análisis). Por consiguiente, $\text{SO}(2)$ está formado por las matrices de la forma

$$M = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

En \mathbf{R}^2 , provisto de la estructura euclídea orientada canónicamente, una matriz de este tipo representa una rotación de ángulo θ alrededor del origen. Así $\text{SO}(2)$ es el grupo de las rotaciones de ángulo arbitrario alrededor del origen.

Segundo caso: $ad - bc = -1$. Las matrices correspondientes son los elementos de $\text{O}(2) \setminus \text{SO}(2)$. Razonemos como antes: a y b no son ambos nulos. De $a^2 + b^2 = bc - ad = 1$, se obtiene

$$a(a + d) + b(b - c) = 0,$$

de donde se sigue la existencia de un $\lambda \in \mathbf{R}$ tal que $d = -a + \lambda b$, $c = b + \lambda a$. Llevando estos valores a $ac + bd = 0$, obtenemos $\lambda(a^2 + b^2) = 0$, luego $\lambda = 0$. El grupo $\text{O}(2)$ está formado por las matrices de la forma

$$\begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \text{ en donde } a^2 + b^2 = 1.$$

Si hacemos $a = \cos \theta$, $b = \sin \theta$, vemos que $\text{O}(2) \setminus \text{SO}(2)$ está formado por las matrices de la forma: $M = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}$. Una matriz de esta forma representa una simetría respecto de la recta de ángulo polar $\frac{\theta}{2}$. Los elementos de $\text{O}(2) \setminus \text{SO}(2)$ son exclusivamente las simetrías respecto de las rectas que pasan por el origen.

Cambios de base

XII.7.4 Sea $(e_i)_{1 \leq i \leq n}$ una base ortonormal del espacio euclídeo E_n . Para que un endomorfismo φ de E_n sea un **automorfismo ortogonal**, es necesario

|| y suficiente que los n vectores $f_i = \varphi(e_i)$ ($1 \leq i \leq n$) formen una base **ortonormal** de E_n .

Demostración. Según lo que antecede, φ es un automorfismo ortogonal si, y sólo si, se tiene:

$$(7) \quad |\varphi(x)|^2 = |x|^2 \quad \text{para todo } x \in E_n.$$

Si hacemos $x = \sum_i x_i e_i$, se tiene, por otra parte:

$$|x|^2 = \sum_i (x_i)^2 \quad \text{y} \quad |\varphi(x)|^2 = \left| \sum_i x_i f_i \right|^2 = \sum_{i=1}^n \sum_{j=1}^n f_i \cdot f_j x_i x_j.$$

La condición (7) se manifiesta entonces por medio de las relaciones $f_i \cdot f_j = \delta_{ij}$ ($i, j = 1, 2, \dots, n$) que expresan que las f_i forman una base ortonormal. c.q.d.

COROLARIO

|| Para que n vectores f_j ($1 \leq i \leq n$) de E_n formen una base **ortonormal** es necesario y suficiente que su matriz, respecto a una base ortonormal (e_i) , sea **ortogonal**.

(Se puede comprobar además que las relaciones (4) expresan el hecho de que los vectores columna de la matriz $M = [a_{ij}]$ forman una base ortonormal.)

Orientación

Sea E un espacio vectorial de dimensión finita n sobre \mathbf{R} . Dadas dos bases ordenadas (e_1, \dots, e_n) , (f_1, \dots, f_n) de E , se dice que la segunda tiene la **misma orientación** que la primera si la matriz que pasa de las (e_i) a las (f_i) tiene un determinante positivo. Las propiedades de los determinantes prueban que la relación binaria definida de esta manera en el conjunto de las bases de E es una *relación de equivalencia*, que determina *dos clases de equivalencia* ⁽¹⁾. **Orientar** E significa elegir una de ambas clases, a cuyos elementos se les llama **bases directas**.

Si E tiene una estructura euclídea, podemos limitarnos a considerar bases ortonormales. Dos bases ortonormales tienen la misma orientación si el determinante de la matriz de cambio es igual a $+1$. Luego *las isometrías que conservan la orientación del espacio euclídeo E_n son las isometrías directas*.

La *Geometría euclídea* constituye el estudio de las propiedades invariantes frente al grupo de las traslaciones y frente al grupo $O(n, \mathbf{R})$.

⁽¹⁾ Esta cuestión será considerada de nuevo en el tomo 3 (*Geometría*).

La *Geometría euclídea orientada* constituye el estudio de las propiedades invariantes frente al grupo de las traslaciones y frente al grupo $SO(n, \mathbf{R})$.

Producto mixto y producto vectorial

DEFINICIÓN XII.7.5

*El **producto mixto** de n vectores V_1, \dots, V_n del espacio euclídeo orientado E_n , es el determinante de las componentes v_{ij} de las V_i en una base ortonormal directa (e_1, \dots, e_n) de E_n .*

Esta definición es independiente de la elección de la base, ya que, si (v'_{ij}) son las componentes de V_1, \dots, V_n en otra base ortonormal directa (e'_i) , se tiene:

$$\det([v'_{ij}]) \det(P) = \det([v_{ij}]),$$

si designamos por P la matriz de cambio de las (e_i) en las (e'_i) , de donde, si

$$P \in SO(n, \mathbf{R}), \quad \det([v'_{ij}]) = \det([v_{ij}]).$$

El producto mixto de V_1, \dots, V_n lo designaremos por (V_1, V_2, \dots, V_n) o por $[V_1, V_2, \dots, V_n]$; es una función n -lineal alternada sobre el espacio vectorial E_n .

TEOREMA XII.7.5

Dados $n - 1$ vectores V_1, \dots, V_{n-1} del espacio euclídeo orientado E_n , existe un vector W único tal que, para todo vector V de E_n , se verifica:

$$V \cdot W = (V, V_1, V_2, \dots, V_{n-1}).$$

*Este vector, llamado **producto vectorial** de los $n - 1$ vectores V_1, \dots, V_{n-1} , se designa por $V_1 \wedge V_2 \wedge \dots \wedge V_{n-1}$. Se reemplaza por su opuesto si se cambia la orientación de E_n .*

En efecto, la aplicación $\varphi: V \mapsto (V, V_1, \dots, V_{n-1})$ es lineal de E_n en \mathbf{R} , y a esta forma lineal φ se le puede asociar, por lo tanto (T. XII.3.2), un vector W único tal que $\varphi(V) = V \cdot W$. c.q.d.

Las componentes W_i del producto vectorial $W = V_1 \wedge \dots \wedge V_{n-1}$ vienen dadas por

$$W_i = (-1)^{i+1} \Delta_i,$$

en donde Δ_i designa el determinante de orden $n - 1$ obtenido suprimiendo la fila i en la matriz de las componentes de los V_k en una base ortonormal.

El producto vectorial es una función $(n - 1)$ -lineal y *alternada* sobre el espacio E_n .

El producto mixto y el producto vectorial son dos ejemplos de propiedades invariantes para $SO(n, \mathbf{R})$, pero no para $O(n, \mathbf{R})$ (ya que las transformaciones ortogonales de determinante -1 cambian el producto mixto y el producto vectorial en sus opuestos).

Capítulo XIII

Formas hermíticas. Teoría espectral Isometrías

En el estudio de las formas hermíticas, que sigue a continuación, seremos mucho más breves que en el capítulo XII, ya que muchas de las demostraciones y de las nociones se inspiran en los mismos principios.

§ XIII.1 GENERALIDADES

- El cuerpo base es \mathbf{C} , cuerpo de los números complejos.

DEFINICIÓN XIII.1.1

Una forma **sesquilineal** definida en el \mathbf{C} -espacio vectorial E es una aplicación $h : E \times E \rightarrow \mathbf{C}$ tal que, para todo $x \in E$, todo $y \in E$, todo $z \in E$, y todo $\lambda \in \mathbf{C}$, se verifica:

$$\begin{aligned} (1) \quad & h(x+y, z) = h(x, z) + h(y, z), \quad h(x, y+z) = h(x, y) + h(x, z), \\ & h(\lambda x, y) = \lambda h(x, y), \\ & h(x, \lambda y) = \bar{\lambda} h(x, y). \end{aligned}$$

DEFINICIÓN XIII.1.2

Una forma **hermítica** definida en el \mathbf{C} -espacio vectorial E es una forma sesquilineal h , sujeta a la condición siguiente:

$$(2) \quad \text{para todo } x \in E \text{ y } y \in E, \quad h(y, x) = \overline{h(x, y)}.$$

(1) y (2) implican:

$$h(x + y, x + y) = h(x, x) + h(y, y) + 2 \operatorname{Re} [h(x, y)] .$$

($\operatorname{Re}(z)$ e $\operatorname{Im}(z)$ designan, respectivamente, la parte real e imaginaria del número complejo z .)

Ejemplos de formas hermiticas

1) En $E = \mathbf{C}^n$, $h(x, y) = \sum_i x_i \bar{y}_i$ con ($x = (x_i)$, $y = (y_i)$).

2) Sea E el espacio de las funciones continuas $f: [0, 1] \rightarrow \mathbf{C}$. La aplicación $(f, g) \mapsto \int_0^1 f(t) \bar{g}(t) dt$ es una forma hermitica.

3) E es el espacio de las sucesiones (u_n) de los números complejos tales que

$$\sum_{n=0}^{\infty} |u_n|^2 < +\infty .$$

Si $u = (u_n)$ y $v = (v_n)$ son elementos de E , veremos (tomo 2) que la serie $\sum_{n=0}^{\infty} u_n \bar{v}_n$ converge. La aplicación $(u, v) \mapsto \sum_{n=0}^{\infty} u_n \bar{v}_n$ es una forma hermitica sobre E .

Matriz de una forma hermitica (en dimensión finita)

Sea h una forma hermitica definida en el \mathbf{C} -espacio vectorial E , de dimensión n . Para toda base (e_1, \dots, e_n) , a la matriz cuadrada

$$H = [h(e_i, e_j)]$$

se le llama *matriz H de h en esta base*.

Sean $x = \sum x_i e_i$ e $y = \sum y_i e_i$, y designemos respectivamente por \mathcal{X} e \mathcal{Y} a las matrices columna de los (x_i) y de los (y_i) . Por un método análogo al del capítulo XII, § 1, se demuestra

$$(3) \quad h(x, y) = {}^t \mathcal{X} H \bar{\mathcal{Y}} ,$$

o sea, si hacemos $h(e_i, e_j) = h_{ij}$:

$$h(x, y) = \sum_{i=1}^n \sum_{j=1}^n h_{ij} x_i \bar{y}_j .$$

Además, la matriz H de h en (e_1, \dots, e_n) verifica, en virtud de (2), la relación

$$(4) \quad {}^tH = \bar{H}.$$

Recíprocamente, si H verifica (4), la fórmula (3) define sobre E una forma hermitica, cuya matriz en la base (e_i) es H .

DEFINICIÓN XIII.1.3

$\left\{ \begin{array}{l} \text{A una } (n, n)\text{-matriz } H \text{ con elementos complejos se le llama } \mathbf{hermitica} \text{ si} \\ {}^tH = \bar{H}. \text{ En otras palabras, la matriz } [h_{ij}] \text{ es hermitica si, cualesquiera} \\ \text{que sean } i, j = 1, 2, \dots, n, \text{ se tiene } h_{ji} = \bar{h}_{ij}. \end{array} \right.$

Cambio de base

Conservemos las notaciones de la fórmula (3). Si (f_1, \dots, f_n) es una nueva base de E , deducida de (e_1, \dots, e_n) por la matriz de cambio P , se demuestra, como en el capítulo XII, § 1, que la matriz H_1 de h en la base (f_1, \dots, f_n) es:

$$(5) \quad H_1 = {}^tPH\bar{P}.$$

Nota. $x \rightarrow h(x, x)$ no es un polinomio respecto a los x_i ; además, puesto que $h(x, x) = \overline{h(x, x)}$ vemos que, para todo $x \in E$, $h(x, x) \in \mathbf{R}$.

Sin embargo, dar la aplicación $x \mapsto h(x, x)$ determina totalmente a la forma $h(x, y)$ en virtud de la fórmula

$$\begin{aligned} 4 h(x, y) &= h(x + y, x + y) - h(x - y, x - y) + \\ &\quad + ih(x + iy, x + iy) - ih(x - iy, x - iy). \end{aligned}$$

Dualidad

Precisaremos de la noción de aplicación semilineal:

DEFINICIÓN XIII.1.4

$\left\{ \begin{array}{l} \text{Si } E \text{ y } F \text{ designan dos } \mathbf{C}\text{-espacios vectoriales, una aplicación } \varphi: E \rightarrow F \\ \text{se llama } \mathbf{semilineal} \text{ si verifica} \end{array} \right.$

$$\forall x_1, x_2 \in E, \quad \varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2) \quad (\text{aditividad})$$

$$\forall \lambda \in \mathbf{C}, \forall x \in E, \quad \varphi(\lambda x) = \bar{\lambda} \varphi(x) \quad (\text{semilinealidad}).$$

Dotado de su ley de grupo abeliano, y de la ley externa $(\lambda, x) \mapsto \bar{\lambda}.x$, F se convierte en un \mathbf{C} -espacio vectorial, que designaremos por \bar{F} . La aplicación idéntica $F \rightarrow F$ no es un isomorfismo de F en \bar{F} .

Dicho esto, afirmar que la aplicación $\varphi : E \rightarrow F$ es semilineal, es lo mismo que enunciar que la aplicación $\varphi : E \rightarrow \bar{F}$, que toma los mismos valores que φ , es lineal.

Podremos pues hablar de la matriz de φ relativa a una elección de bases en E y F (ya que dar una base de F equivale a dar una base de \bar{F}). c.q.d.

Sea h una forma hermítica definida en el \mathbf{C} -espacio E . Para todo $x \in E$, la aplicación $h(., y) : x \mapsto h(x, y)$ es una forma lineal sobre E , en virtud de (1).

Y, siempre en virtud de (1), la aplicación $J : y \mapsto h(., y)$ es *semilineal* de E en E^* . Si E es de dimensión finita n , la matriz de J en una base (e_1, \dots, e_n) de E y en su base dual (e_1^*, \dots, e_n^*) es precisamente la matriz de h en (e_1, \dots, e_n) . El rango de esta matriz es, por lo tanto, independiente de la base elegida (lo cual se deduce también de (5)).

DEFINICIÓN XIII.1.5

Si E designa un \mathbf{C} -espacio vectorial de dimensión finita n , el **rango** de una forma hermítica h definida en E es el de la aplicación $J : E \rightarrow E^*$, tal que

$$J(y) = h(., y)$$

Se dice que h es **no degenerada** si su rango es máximo, por lo tanto igual a $\dim(E)$. Se le llama **degenerada** en caso contrario.

El rango de h es el de su matriz en una base cualquiera. Una forma no degenerada se halla, entonces, caracterizada por el hecho de que esta matriz es regular. La aplicación J es entonces una *biyección*. Dada su gran importancia, enunciamos el resultado obtenido de la forma siguiente:

TEOREMA XIII.1.1

Sea h una forma hermítica no degenerada sobre el \mathbf{C} -espacio vectorial E de dimensión finita. Para toda forma lineal φ definida en E , existe un elemento $y_\varphi \in E$ y sólo uno tal que, para todo $x \in E$, se verifique

$$\varphi(x) = h(x, y_\varphi).$$

Este resultado generaliza XII.3.2.

Nota. Contrariamente a lo que ocurre con las formas bilineales, cualquiera que sea $x \in E$, la aplicación $h(x, \cdot): y \mapsto h(x, y)$ no es \mathbf{C} -lineal, sino que verifica $h(x, \cdot)(\lambda y) = h(x, \lambda y) = \bar{\lambda} h(x, \cdot)(y)$ para todo $\lambda \in \mathbf{C}$. $h(x, \cdot)$ es, por lo tanto, *semilineal*.

Igualmente a como se ha hecho en el caso de formas cuadráticas, se define la noción de ortogonalidad respecto de una forma hermitica h . Dos elementos $x, y \in E$ son **ortogonales** si $h(x, y) = 0$, y se escribe entonces $x \perp y$. Si A es una parte de E , **el ortogonal** de A (designado por A^\perp) es el conjunto de los $y \in E$ tales que $x \perp y$ para todo $x \in A$. A^\perp es siempre un subespacio de E , y para los detalles remitimos al lector al capítulo XII, § 3.

Señalemos el importante teorema análogo al XII.3.3:

TEOREMA XIII.1.2

Si h es una forma hermitica no degenerada definida en el \mathbf{C} -espacio vectorial E de dimensión finita n , para todo subespacio vectorial F de E , se tiene: $\dim(F) + \dim(F^\perp) = n$, y $(F^\perp)^\perp = F$. En particular, para toda parte A de E , $(A^\perp)^\perp = \text{Vect}(A)$.

La demostración es calcada de la de XII.3.3, utilizando XIII.1.1.

§ XIII.2 CLASIFICACIÓN DE LAS FORMAS HERMÍTICAS SOBRE UN ESPACIO DE DIMENSIÓN FINITA

Se obtiene, sin dificultad, utilizando un método análogo al de XII.3.4, el:

TEOREMA XIII.2.1

Para toda forma hermitica h definida en un \mathbf{C} -espacio vectorial de dimensión finita n , existe una base (e_1, \dots, e_n) **ortogonal**, es decir, que verifica

$$h(e_i, e_j) = 0 \quad \text{para } i \neq j.$$

Una base ortogonal (e_i) se caracteriza por el hecho de que la matriz de h en esta base es diagonal. Se tiene pues (con una numeración conveniente de la base):

$$(1) \quad h\left(\sum x_i e_i, \sum y_j e_j\right) = \sum_{i=1}^r \alpha_i x_i \bar{y}_i, \quad \text{con } \alpha_i = h(e_i, e_i), \quad \alpha_1 \neq 0, \dots, \alpha_r \neq 0,$$

en donde r designa el *rango* de h .

Observemos que todos los α_i son reales, pues $h(e_i, e_i) = \overline{h(e_i, e_i)}$ según la definición XIII.1.2.

Reemplazando, si es necesario, e_i por $\frac{e_i}{\sqrt{|a_i|}}$, la fórmula (1) se convierte en:

$$(2) \quad h(\sum x_i e_i, \sum y_j e_j) = x_1 \bar{y}_1 + \cdots + x_p \bar{y}_p - x_{p+1} \bar{y}_{p+1} - \cdots - x_r \bar{y}_r.$$

El número p no depende de la base (e_i) elegida, y la demostración es idéntica a la de XII.4.3, con la diferencia de que $Q(x)$ se reemplaza por $h(x, x)$. Con otras palabras, *la ley de inercia de Sylvester* es válida para las formas hermiticas. Estas quedan clasificadas por los pares de enteros $(p, r - p)$, en donde $0 \leq p \leq r \leq n$, y en donde r es *el rango*. Al par $(p, r - p)$ se le llama *signatura* o *tipo* de la forma h .

DEFINICIÓN XIII.2.1

$\left\{ \begin{array}{l} \text{A una forma hermitica } h \text{ definida en el } \mathbf{C}\text{-espacio vectorial } E, \text{ se le llama} \\ \text{positiva si} \\ h(x, x) \geq 0 \text{ para } x \in E; \\ \text{y definida positiva si } x \in E \text{ y } x \neq 0 \text{ implican } h(x, x) > 0. \end{array} \right.$

Si E es de dimensión finita n , h es *positiva* si es del tipo $(r, 0)$, puesto que en una base conveniente se reduce a $\sum_{i=1}^r x_i \bar{y}_i$. Es *definida positiva* si es del tipo $(n, 0)$ puesto que en una base conveniente se reduce a $\sum_{i=1}^n x_i \bar{y}_i$.

El estudio de las formas negativas (resp. definidas negativas) se reduce al de las formas positivas (resp. definidas positivas), si se toma la forma opuesta $-h$.

§ XIII.3 ESPACIOS PREHILBERTIANOS DE DIMENSIÓN FINITA

DEFINICIÓN XIII.3.1

$\left\{ \begin{array}{l} \text{Un espacio prehilbertiano real [resp. complejo] es un espacio vectorial} \\ \text{sobre } \mathbf{R} \text{ [resp. sobre } \mathbf{C}] \text{ provisto de una forma bilineal simétrica [resp. her-} \\ \text{mítica] definida positiva, llamada producto escalar.} \end{array} \right.$

Los espacios prehilbertianos reales de dimensión finita han sido estudiados en el capítulo XII con el nombre de *espacios euclídeos*.

En lo que sigue, únicamente consideraremos espacios prehilbertianos sobre \mathbf{C} . A cada propiedad de los espacios prehilbertianos complejos corresponde una propiedad de los espacios prehilbertianos reales, por restricción del cuerpo de escalares.

En general, el producto escalar de un espacio prehilbertiano E se designa por $(x | y)$. Se escribe: $(x | x) = |x|^2$.

Si E es de dimensión finita n , y si (e_1, \dots, e_n) es una base ortonormal, se tiene:

$$(1) \quad \left| \sum_i x_i e_i \right|^2 = |x_1|^2 + |x_2|^2 + \dots + |x_n|^2.$$

Por otro lado, si $x = \sum_i x_i e_i$, $(x | e_i) = x_i$, y la fórmula (1) nos da:

$$(2) \quad |x|^2 = \sum_i |(x | e_i)|^2 \quad (\text{fórmula de Parseval}).$$

Finalmente, aplicando XIII.1.1, vemos que toda forma lineal φ definida en E se escribe de una manera y sólo una, en la forma

$$\varphi(x) = (x | y),$$

en donde y designa un elemento de E .

Métrica asociada a un espacio prehilbertiano

Primeramente demostraremos la *desigualdad de Cauchy-Schwarz*:

TEOREMA XIII.3.1

$$\left\| \begin{array}{l} \text{Para todo par de elementos } x, y \text{ de un espacio prehilbertiano } E, \text{ se tiene } (1): \\ (3) \quad |(x | y)|^2 \leq |x|^2 \cdot |y|^2; \\ y, \text{ si } y \neq 0, \text{ la desigualdad (3) se transforma en igualdad si, y sólo si,} \\ \text{existe } \lambda \in \mathbf{C} \text{ tal que } x + \lambda y = 0. \end{array} \right.$$

Demostración. Si $(x | y) = 0$, (3) es evidente. Si $(x | y) \neq 0$, x e y no son ambos nulos. Tenemos que para todo $\lambda \in \mathbf{C}$, se verifica

$$(x + \lambda y | x + \lambda y) \geq 0.$$

Se obtiene:

$$|x|^2 + \bar{\lambda}(x | y) + \lambda \overline{(x | y)} + |\lambda|^2 |y|^2 \geq 0, \quad \text{con } |y| > 0.$$

(1) La desigualdad (1) es válida para un espacio vectorial sobre \mathbf{R} (resp. sobre \mathbf{C}) provisto de una forma cuadrática (resp. hermitica) positiva, no necesariamente definida; pero la igualdad puede presentarse, en cambio, sin que los vectores x, y sean colineales (ver p. 464).

En esta relación, sustituimos λ por $\rho \frac{(x|y)}{|(x|y)|}$, en donde $\rho \in \mathbf{R}$. Se obtiene:

$$(4) \quad |x|^2 + 2\rho |(x|y)| + \rho^2 |y|^2 \geq 0.$$

El trinomio del primer miembro de (4) es ≥ 0 para todo $\rho \in \mathbf{R}$, por lo tanto su discriminante es ≤ 0 , de donde $|(x|y)|^2 \leq |x|^2 |y|^2$, y (3) queda establecida.

Si $x + \lambda y = 0$ ($\lambda \in \mathbf{C}$), (3) es una igualdad. Recíprocamente, si (3) es una igualdad, según (4), existe un $\rho \in \mathbf{R}$ tal que $(|x| + \rho |y|)^2 = 0$, luego un $\lambda \in \mathbf{C}$ tal que $(x + \lambda y | x + \lambda y) = 0$, lo cual exige $x + \lambda y = 0$. c.q.d.

COROLARIO

En todo espacio prehilbertiano E , la aplicación $N: E \rightarrow \mathbf{R}_+$, con $N(x) = (x|x)^{1/2}$, es una norma. Además, para que se verifique

$$N(x + y) = N(x) + N(y),$$

es necesario y suficiente (si $y \neq 0$) que exista un $\lambda \in \mathbf{R}_-$ tal que $x + \lambda y = 0$.

Demostración

a) Para ver que N es una norma, es suficiente establecer la desigualdad triangular: $N(x + y) \leq N(x) + N(y)$.

Entonces se tiene:

$$\begin{aligned} N^2(x + y) &= (x + y | x + y) = (x|x) + (y|y) + 2 \operatorname{Re} (x|y) \\ &= N^2(x) + N^2(y) + 2 \operatorname{Re} (x|y) \leq N^2(x) + N^2(y) + \\ &\quad + 2 |(x|y)| \leq N^2(x) + N^2(y) + 2 N(x) N(y) \\ &= (N(x) + N(y))^2 \quad (\text{según (3)}), \end{aligned}$$

de donde

$$N(x + y) \leq N(x) + N(y).$$

b) Si $x + \lambda y = 0$, $\lambda \in \mathbf{R}_-$, $N(x + y) = N(x) + N(y)$. Recíprocamente, si $N(x + y) = N(x) + N(y)$, esto exige $\operatorname{Re} [(x|y)] = |x| \cdot |y|$. Puesto que $|(x|y)| \leq |x| \cdot |y|$, se tiene: $(x|y) = |x| \cdot |y|$. Volviendo a considerar los cálculos de XIII.3.1, vemos que el valor de λ tal que $(x + \lambda y | x + \lambda y) = 0$ es entonces real y negativo, puesto que

$$\lambda = \rho \frac{(x|y)}{|(x|y)|} = \rho = - \frac{|x|}{|y|} \leq 0.$$

De donde $x + \lambda y = 0$, con $\lambda \in \mathbf{R}_-$. c.q.d.

Nota. Si E es de dimensión finita n , pasando a las coordenadas en una base ortonormal, (3) se transforma en

$$(5) \quad \left| \sum_{i=1}^n x_i \bar{y}_i \right|^2 \leq \left(\sum_{i=1}^n |x_i|^2 \right) \left(\sum_{i=1}^n |y_i|^2 \right).$$

Es posible demostrar directamente (5).]

Del mismo modo que en el capítulo XII, § 5, se define la *suma directa de un número finito de espacios prehilbertianos*, y la *estructura prehilbertiana inducida en un subespacio* de un espacio prehilbertiano.

Nos contentaremos con enunciar los teoremas análogos a los del capítulo XII, § 5, sin demostración, ya que permanecen inalterados.

TEOREMA XIII.3.2

|| Sea E un espacio prehilbertiano de dimensión finita. Para todo subespacio F de E , E es suma directa de F y F^\perp .

A F^\perp se le llama **suplementario ortogonal** de F .

TEOREMA XIII.3.3

|| Sea E un espacio prehilbertiano suma directa de los subespacios F_1, \dots, F_k ortogonales dos a dos. Si $x = \sum_{i=1}^k x_i$ es la descomposición de $x \in E$ sobre los F_i , se tiene:

$$(6) \quad |x|^2 = \sum_i |x_i|^2,$$

|| e.d. E es isomorfo a la **suma directa de los espacios prehilbertianos F_i** .

Caso particular. Si E es de dimensión finita n , y si (e_1, \dots, e_n) es una base ortonormal, (6) se reduce a la *fórmula de Parseval*:

$$|x|^2 = \sum_{i=1}^n |(x | e_i)|^2.$$

TEOREMA XIII.3.4

|| Todo sistema ortonormal de un espacio prehilbertiano de dimensión finita se puede completar hasta obtener una base ortonormal del espacio.

TEOREMA XIII.3.5

Sea (v_1, \dots, v_n) una base cualquiera de un espacio prehilbertiano E .
 Existe una base ortonormal **única** de E , a saber (e_1, \dots, e_n) , que verifica las condiciones siguientes;

para todo entero p , $(e_p | v_p) > 0$;
 para todo entero p , $\text{Vect}(e_1, \dots, e_p) = \text{Vect}(v_1, \dots, v_p)$.

Teniendo en cuenta las definiciones del § 5, se puede enunciar XIII.3.5 en la forma equivalente siguiente: para toda matriz invertible M , existe una matriz unitaria U y una matriz triangular superior T tales que $M = UT$.

Extensión de la desigualdad de Cauchy-Schwarz

Análogamente a XII.5.7, se tiene:

XIII.3.6 Sea h una forma hermitica **positiva** definida en un \mathbf{C} -espacio vectorial E . Entonces

$$(\forall x \in E) (\forall y \in E) \quad |h(x, y)|^2 \leq h(x, x) h(y, y).$$

Demostración. Si $(x | y) = 0$, la desigualdad es evidente. Si no, en la desigualdad $h(x + \lambda y, x + \lambda y) \geq 0$ (válida para todo $\lambda \in \mathbf{C}$) reemplazamos λ por $\rho \frac{h(x, y)}{|h(x, y)|}$, en donde $\rho \in \mathbf{R}$. Se obtiene:

$$(7) \quad h(x, x) + 2\rho |h(x, y)| + \rho^2 h(y, y) \geq 0.$$

Si $h(y, y) > 0$, el primer miembro de (7) es un trinomio en ρ , de signo constante, de donde $|h(x, y)|^2 - h(x, x) h(y, y) \leq 0$.

Si $h(y, y) = 0$, el primer miembro de (7) es una función afín en ρ , de signo constante, por lo tanto constante, luego $h(x, y) = 0$.]

Convenimos en establecer que una forma hermitica h sobre un \mathbf{C} -espacio vectorial E es *no degenerada* si la aplicación

$$J : E \rightarrow E^*, \quad y \mapsto J(., y)$$

(en donde $\forall x \in E$, $J(., y)(x) = J(x, y)$) es *inyectiva*. Así como de XII.5.7 se deduce el corolario de XII.5.7, así también de XIII.3.6 se deduce :

COROLARIO

|| Sea h una forma hermítica **positiva** en un \mathbf{C} -espacio vectorial. Para que h sea **no degenerada**, es necesario y suficiente que sea **definida positiva**

§ XIII.4 PROYECCIONES Y SIMETRÍAS

● Ahora consideraremos un espacio prehilbertiano E de dimensión finita n_i , cuyo producto escalar se designará por $(x | y)$, la norma del vector $x \in E$ se designará por $|x|$, y la distancia $|x - y|$ de los puntos $x, y \in E$ se podrá indicar por $d(x, y)$.

TEOREMA XIII.4.1

|| Sean H un subespacio afín de E , y a un punto cualquiera de E . Existe un punto $q(a)$, y sólo uno, tal que $q(a) \in H$, y que:

$$d(a, H) = d(a, q(a)).$$

$A(q(a))$ se le llama **proyección ortogonal de a sobre H** .

Este punto es el único punto $b \in H$ tal que $b - a$ es ortogonal a la dirección H_0 de H .

Finalmente, la aplicación $a \mapsto q(a)$ es una **aplicación afín, epiyectiva, llamada proyección ortogonal sobre H** .

Demostración. Calcada de la de XII.6.1, reemplazando los productos escalares euclídeos por los productos escalares hermíticos. ||

Como en el capítulo XII, § 6, se define la simetría respecto a un subespacio afín H de E , que es una biyección afín de E en E , involutiva.

Conjunto de los puntos equidistantes de dos puntos de E

Los resultados del capítulo XII, § 6, no se generalizan.

Sean $a, b \in E$ dos puntos distintos. Busquemos el conjunto de los $x \in E$ tales que

$$d(x, a) = d(x, b).$$

Esta relación se escribe:

$$(x - a | x - a) = (x - b | x - b),$$

o sea, desarrollando:

$$2 \operatorname{Re} [(x | b)] - 2 \operatorname{Re} [(x | a)] = |b|^2 - |a|^2.$$

$$(1) \quad 2 \operatorname{Re} [(x | b - a)] = |b|^2 - |a|^2.$$

La aplicación $\rho : x \mapsto \operatorname{Re} [(x | b - a)]$ no es \mathbf{C} -lineal. Sin embargo, E se convierte en un \mathbf{R} -espacio vectorial, de dimensión $2n$, por *restricción de los escalares a \mathbf{R}* . Designamos por $E_{(\mathbf{R})}$ este espacio vectorial, y ρ es entonces una forma lineal sobre $E_{(\mathbf{R})}$. El conjunto \mathcal{H} de los puntos x que verifican (1) es, entonces, un *hiperplano afín de $E_{(\mathbf{R})}$* , por lo tanto un espacio vectorial de dimensión $2n - 1$ sobre \mathbf{R} . Puesto que \mathcal{H} es de dimensión impar sobre \mathbf{R} , *no es* un subespacio vectorial de E sobre \mathbf{C} .

El ortogonal H_0 de $(b - a)$ en E es un hiperplano vectorial de E . Si m es un punto cualquiera de \mathcal{H} , es claro que el hiperplano afín $m + H_0 = H$ está contenido en \mathcal{H} . Entonces podemos decir que el par (a, b) admite una *infinitud de hiperplanos medios*, todos ellos contenidos en \mathcal{H} , y paralelos a H_0 .

§ XIII.5 GRUPO UNITARIO

Sea h una forma hermitica cualquiera sobre el \mathbf{C} -espacio vectorial E . Se dice que un automorfismo φ de E es **unitario** (para la forma h) si verifica

$$(1) \quad \forall x \in E, \forall y \in E \quad h(\varphi(x), \varphi(y)) = h(x, y).$$

Lo que equivale a:

$$\forall x \in E \quad h(\varphi(x), \varphi(x)) = h(\varphi(y), \varphi(y)).$$

Es preciso observar que cuando E es de dimensión finita y h es no degenerado, si $\varphi \in \mathcal{L}(E)$, (1) implica $\varphi \in \operatorname{GL}(E)$ (luego φ es unitario). Los automorfismos de E que verifican (1) constituyen un subgrupo de $\operatorname{GL}(E)$, llamado **grupo unitario** asociado a la forma h , y designado $U(h)$.

Si supiésemos clasificar las formas hermiticas sobre E (con otras palabras: si se conociesen las condiciones necesarias y suficientes para que dos formas h_1 y h_2 fuesen equivalentes, o para que se expresasen, en bases convenientes, por medio de la misma matriz, cf. Cap. XII, principio § 4), sabríamos también clasificar los grupos unitarios sobre E .

Caso en que E es de dimensión finita: grupo especial unitario de una forma hermitica

Sea h una forma hermitica *no degenerada* definida en el \mathbf{C} -espacio vectorial E de dimensión n . Es fácil caracterizar los elementos de $U(h)$ por su matriz en una base. Se tiene:

XIII.5.1 Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de E , y H la matriz de h en \mathcal{B} . Para que un endomorfismo $u \in \mathcal{L}(E)$ sea un elemento de $U(h)$, es necesario y suficiente que su matriz M_u en \mathcal{B} verifique

$${}^t M_u H \overline{M_u} = H.$$

En consecuencia, el conjunto Γ_H de las matrices $M \in M_n(\mathbf{C})$ que verifican ${}^t M H \overline{M} = H$ forma un subgrupo de $GL_n(\mathbf{C})$, y la aplicación $U(h) \rightarrow \Gamma_H$, $u \mapsto M_u$ es un isomorfismo de grupos.

La demostración es análoga a la de XII.7.1.

XIII.5.1 es interesante cuando H es diagonal (e.d. cuando \mathcal{B} es ortogonal). Particularizando más, cuando \mathcal{B} es ortonormal (lo que ocurrirá únicamente cuando h sea definida positiva o negativa), se tiene $H = I_n$, y al grupo Γ_H se le llama entonces grupo de las matrices unitarias de orden n sobre \mathbf{C} . Este caso se estudiará más adelante.

Ejemplo

Si \mathcal{B} es ortogonal, y si

$$M_u = \begin{bmatrix} \xi_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \xi_n \end{bmatrix} \quad \text{con} \quad \forall i \quad |\xi_i| = 1,$$

es claro que $M_u \in \Gamma_H$, por lo tanto $u \in U(h)$.

De la relación ${}^t M_u H \overline{M_u} = H$, se deduce

$$|\det(M_u)|^2 = 1 \quad (\text{pues } \det(H) \neq 0).$$

De donde:

XIII.5.2 La aplicación $u \mapsto \det(u)$ es un homomorfismo **epiyectivo** de $U(h)$ en el grupo U de los números complejos de módulo 1.

Demostración (resumida). El único punto no evidente es la epiyectividad. A este fin sea $\mathcal{B} = (e_1, \dots, e_n)$ una base ortogonal para h , y sea $\zeta \in U$. El endomorfismo u de E tal que $u(e_1) = \zeta e_1$ y $\forall i \geq 2 \quad u(e_i) = e_i$ pertenece a $U(h)$, y verifica $\det(u) = \zeta$.]

Se da la siguiente:

DEFINICIÓN XIII.5.1

Se llama **grupo unitario especial de h** , y se designa por $SU(h)$, al subgrupo de $U(h)$ formado por los $u \in U(h)$ tales que $\det(u) = 1$.

Según XIII.5.2, $SU(h)$ es un subgrupo normal de $U(h)$, y se tiene un isomorfismo canónico de $U(h)/SU(h)$ en U .

Observemos finalmente que si dos formas hermiticas son equivalentes (e.d. tienen la misma signatura), sus grupos unitarios (resp. unitarios especiales) son isomorfos.

Grupo unitario de una forma hermitica definida positiva

En particular, sea h una forma hermitica definida positiva sobre un espacio vectorial E de dimension n . En una base ortonormal (e_1, \dots, e_n) se tendrá

$$h(x, y) = \sum_{i=1}^n x_i \bar{y}_i \quad (\text{si } x = \sum x_i e_i \text{ y } y = \sum y_j e_j).$$

Todos los grupos unitarios así obtenidos son entonces isomorfos al grupo unitario asociado a la forma hermitica $\sum x_i \bar{y}_i$ definida en \mathbf{C}^n .

DEFINICIÓN XIII.5.2

Al grupo de los automorfismos del \mathbf{C} -espacio vectorial \mathbf{C}^n que dejan invariante la forma hermitica:

$$(2) \quad (x, y) \mapsto \sum_{i=1}^n x_i \bar{y}_i$$

se le llama **grupo unitario de orden n** , y se designa por $U(n, \mathbf{C})$ o $U(n)$.

Con las notaciones de la definición XIII.5.1, sean M la matriz de un endomorfismo φ en la base canónica de \mathbf{C}^n , y \mathcal{X} e \mathcal{Y} las matrices columna de las coordenadas de $x, y \in \mathbf{C}^n$, y \mathcal{X}_1 e \mathcal{Y}_1 las de $x_1 = \varphi(x)$, $y_1 = \varphi(y)$. La condición para que φ deje invariante el producto escalar (2) es:

$$\forall \mathcal{X}, \forall \mathcal{Y}, \quad {}^t \mathcal{X}_1 \bar{\mathcal{Y}}_1 = {}^t \mathcal{X} \bar{\mathcal{Y}} = {}^t \mathcal{X} I_n \bar{\mathcal{Y}},$$

pero sabemos que $\mathcal{X}_1 = M\mathcal{X}$, $\mathcal{Y}_1 = M\mathcal{Y}$, por lo que la condición anterior se convierte entonces en:

$$\forall \mathcal{X}, \forall \mathcal{Y} : \quad {}^t \mathcal{X} {}^t M \bar{M} \bar{\mathcal{Y}} = {}^t \mathcal{X} I_n \bar{\mathcal{Y}},$$

lo que equivale a:

$$(3) \quad {}^t M \overline{M} = I_n.$$

XIII.5.3 El grupo unitario de orden n : $U(n, \mathbf{C})$, es canónicamente isomorfo al grupo multiplicativo de las matrices M que verifican:

$${}^t M \overline{M} = I_n,$$

(a estas matrices se les llama **unitarias**).

En efecto, la relación (3) implica $|\det(M)| = 1$, por lo tanto M es invertible. La proposición se sigue entonces del estudio precedente. \square

DEFINICIÓN XIII.5.3

$\left\{ \begin{array}{l} \text{Para toda matriz cuadrada } M \in M_n(\mathbf{C}), \text{ se llama } \textbf{adjunta} \text{ de } M \text{ a la} \\ \text{matriz } {}^t \overline{M}, \text{ y se designa por } M^*. \end{array} \right.$

Se tiene: $(M^*)^* = M$.

La aplicación $M \mapsto M^*$ es aditiva, pero no es \mathbf{C} -lineal sobre $M_n(\mathbf{C})$. Por restricción de los escalares a \mathbf{R} , $M_n(\mathbf{C})$ se convierte en un \mathbf{R} -espacio vectorial (de dimensión $2n^2$). La aplicación $M \mapsto M^*$ es lineal para esta estructura de \mathbf{R} -espacio vectorial. La fórmula $(\lambda M)^* = \overline{\lambda} M^*$ significa que $M \mapsto M^*$ es *semilineal* para la estructura de \mathbf{C} -espacio vectorial de $M_n(\mathbf{C})$.

Además, se verifica: $(MN)^* = N^* M^*$.

Observemos finalmente el teorema siguiente, cuya demostración es análoga a la de XII.7.3:

XIII.5.4 Sea $(e_i)_{1 \leq i \leq n}$ una base **ortonormal** del espacio prehilbertiano E de dimensión n . Para que un endomorfismo $\varphi \in \mathcal{L}(E)$ sea **unitario** es necesario y suficiente que $(\varphi(e_i))_{1 \leq i \leq n}$ sea una base **ortonormal**.

Ejemplo 1

El grupo $U(1, \mathbf{C})$ es el grupo de los $\zeta \in \mathbf{C}$ tales que $\zeta \overline{\zeta} = 1$, por lo tanto es el grupo U de los números complejos de módulo 1, isomorfo a $SO(2)$.

Consideremos ahora un entero n cualquiera, y sea $M \in U(n, \mathbf{C})$. (3) implica

$$|\det M|^2 = 1, \text{ luego } \det(M) \in U.$$

Con otras palabras, la aplicación $M \mapsto \det(M)$ es un homomorfismo del grupo $U(n, \mathbf{C})$ en el grupo U . Observemos que *este homomorfismo es epiyectivo*, pues si

$$M = \begin{bmatrix} e^{i\theta} & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & 1 \end{bmatrix}, \quad \text{se tiene: } M \in U(n, \mathbf{C}) \quad \text{y} \quad \det(M) = e^{i\theta}.$$

Sea $SU(n, \mathbf{C})$ el núcleo del homomorfismo $M \mapsto \det(M)$. Lo que precede nos demuestra que $SU(n, \mathbf{C})$ es un *subgrupo normal* de $U(n, \mathbf{C})$, y que el grupo cociente $U(n, \mathbf{C})/SU(n, \mathbf{C})$ es isomorfo a $U(1, \mathbf{C})$. Se da la siguiente

DEFINICIÓN XIII.5.4

$\left\{ \begin{array}{l} \text{Se llama } \mathbf{grupo\ especial\ unitario\ de\ orden\ } n, \text{ y se designa por} \\ \text{SU}(n, \mathbf{C}) \text{ (o } \text{SU}(n)) \text{ al grupo de las matrices } M \in U(n, \mathbf{C}) \text{ tales que} \\ \det(M) = 1. \end{array} \right.$

Contrariamente a lo que ocurría con el grupo $SO(n)$, el grupo $U(n)$ no puede servir para definir una nueva noción de orientación para \mathbf{C}^n , que estaría, en este caso, ligada al cuerpo \mathbf{C} . Sin entrar en detalles, señalemos que esto se debe al hecho siguiente: dada una matriz unitaria cualquiera M , es posible encontrar una aplicación *continua* $f: I \rightarrow U(n)$, en donde I es un intervalo de \mathbf{R} , $I = [a, b]$, tal que $f(a) = M$ y $f(b) \in SU(n)$. (En efecto, es suficiente considerar la aplicación $f: [0, \alpha] \rightarrow U(n)$ tal que $f(\theta) = e^{i\theta/n} M$, en donde se ha puesto $\det(M) = e^{-i\alpha}$).

Por el contrario, es evidentemente imposible pasar con continuidad (e.d. por medio de una aplicación continua con valores en $O(n)$) de una matriz ortogonal indirecta a una matriz especial ortogonal. Con otras palabras, $O(n)$ no es *conexo* (cf. tomo 2, Cap. III).

Ejemplo 2

El grupo $SU(1, \mathbf{C})$ se reduce a $\{1\}$, lo mismo que el grupo $SO(1, \mathbf{R})$.

Ejemplo 3

Vamos a determinar el grupo $SU(2, \mathbf{C})$ (lo que determinará el grupo $U(2, \mathbf{C})$, en virtud de la definición XIII.5.3).

Buscamos, pues, las matrices $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tales que

$${}^t M \overline{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad \det(M) = 1,$$

es decir:

$$(4) \quad \begin{cases} a\bar{a} + b\bar{b} = 1 \\ c\bar{c} + d\bar{d} = 1 \\ a\bar{c} + b\bar{d} = 0 \end{cases} \quad (5) \quad ad - bc = 1.$$

Por (5), a y b no son ambos nulos, y de $a\bar{a} + b\bar{b} = ad - bc = 1$, se obtiene:

$$a(\bar{a} - d) + b(\bar{b} + c) = 0,$$

de donde se sigue la existencia de un $\lambda \in \mathbf{C}$ tal que $d = \bar{a} - \lambda b$, $c = -\lambda a - \bar{b}$. Si llevamos estas relaciones a $a\bar{c} + b\bar{d} = 0$, obtendremos:

$$-\bar{\lambda}(a\bar{a} + b\bar{b}) = 0, \quad \text{es decir} \quad \lambda = 0, \quad d = \bar{a} \quad \text{y} \quad c = -\bar{b}.$$

Por lo tanto M , se escribe:

$$M = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}, \quad \text{con} \quad a\bar{a} + b\bar{b} = 1,$$

y el recíproco es inmediato.

El lector deseoso de profundizar en las relaciones notables que ligán los grupos $SU(2, \mathbf{C})$, $SO(3, \mathbf{R})$ y el grupo de los cuaternarios de norma 1, así como el grupo de las homografías $z \mapsto \frac{az + b}{cz + d}$ de \mathbf{C} , podrá consultar con provecho el problema n.º 5.

Para terminar este §, observemos que se tiene una inyección natural

$$O(n) \rightarrow U(n, \mathbf{C}),$$

que envía $SO(n)$ a $SU(n, \mathbf{C})$. Pues toda matriz con elementos reales y ortogonal (resp. especial ortogonal) es evidentemente unitaria (resp. especial unitaria).

§ XIII.6 TEORÍA ESPECTRAL (FORMAS HERMÍTICAS)

● En lo sucesivo aun sin indicarlo explícitamente E designará un espacio vectorial sobre \mathbf{C} , prehilbertiano, de dimensión finita n .

Recordemos (T. XIII.1.1) que, para toda forma lineal u definida en E , existe un elemento $y \in E$ único tal que, para todo $x \in E$, se tiene:

$$u(x) = (x | y).$$

Como aplicación, sea $\varphi \in \mathcal{L}(E)$ un endomorfismo de E , para todo $y \in E$, la aplicación $x \mapsto (\varphi(x) | y)$ es una forma lineal definida en E . Existe, pues, un $y' \in E$ único tal que, para todo $x \in E$, se verifica:

$$(\varphi(x) | y) = (x | y').$$

Es fácil ver que la aplicación $y \mapsto y'$ es lineal de E en E , y ello nos induce a establecer la siguiente:

DEFINICIÓN XIII.6.1

Para todo endomorfismo $\varphi \in \mathcal{L}(E)$, se llama endomorfismo **adjunto** de φ , y se designa por φ^* , al endomorfismo de E definido por

$$(1) \quad \forall x \in E, \forall y \in E \quad (\varphi(x) | y) = (x | \varphi^*(y)).$$

Sea (e_1, \dots, e_n) una base ortonormal de E , si M y N designan las matrices de φ y de φ^* en (e_i) , y \mathcal{X} , \mathcal{Y} las matrices columna de las coordenadas de x e y en (e_i) , (1) equivale a:

$${}^t(M\mathcal{X}) \cdot \overline{\mathcal{Y}} = {}^t\mathcal{X}(\overline{N}\overline{\mathcal{Y}}),$$

$${}^t\mathcal{X} {}^tM\overline{\mathcal{Y}} = {}^t\mathcal{X}\overline{N}\overline{\mathcal{Y}}.$$

Hemos probado así que ${}^tM = \overline{N}$, es decir:

$$(2) \quad N = \overline{{}^tM}.$$

Se verifica que la aplicación $\varphi \mapsto \varphi^*$ es una biyección de $\mathcal{L}(E)$ en $\mathcal{L}(E)$, *semi-lineal e involutiva*, que verifica $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$.

DEFINICIÓN XIII.6.2

Un endomorfismo $\varphi \in \mathcal{L}(E)$ se llama **autoadjunto** o **hermítico** si $\varphi = \varphi^*$, es decir si

$$(3) \quad (\varphi(x) | y) = (x | \varphi(y)) \quad \text{para } x, y \in E.$$

Un endomorfismo autoadjunto, según (2), se halla caracterizado por el hecho de que su matriz M , en toda base ortonormal, es hermítica (${}^tM = \overline{M}$).

TEOREMA XIII.6.1

|| *Un operador autoadjunto $\varphi \in \mathcal{L}(E)$ tiene todos sus valores propios reales, es diagonalizable, y sus subespacios propios son ortogonales dos a dos. En consecuencia, E es suma directa de los subespacios propios de φ .*

Demostración. Razonemos por recurrencia sobre n . El teorema es evidente para $n = 1$. Sean $\lambda_1, \dots, \lambda_p$ los valores propios distintos de φ , y E_1, \dots, E_p los subespacios propios correspondientes. Pongamos $F_1 = E_1^\perp$. Si $F_1^\perp = \{0\}$, $E_1 = E$, y φ es una homotecia. Su matriz en cualquier base es escalar, y puesto que dicha matriz debe ser hermitica en las bases ortonormales, su valor propio λ_1 es necesariamente real, y el teorema está demostrado.

Si $F_1 \neq \{0\}$, se tiene: $\dim(F_1) \leq n - 1$. Probemos que F_1 es estable respecto de φ . Cualesquiera que sean $x \in E_1$ e $y \in F_1$, se tiene:

$$(\varphi(x) | y) = (x | \varphi(y)) = \lambda_1(x | y) = 0,$$

luego $\varphi(y) \in F_1$, o sea $\varphi(F_1) \subset F_1$. Tomando x e y en F_1 , (2) muestra que la restricción φ_1 de φ a F_1 es un operador autoadjunto para la estructura inducida. En virtud de la hipótesis de recurrencia, F_1 es entonces la suma directa de los subespacios propios de φ_1 , que son ortogonales dos a dos. Por lo tanto, E es suma directa de E_1 y de los subespacios propios de φ_1 , y estos últimos se hallan incluidos, a priori, en los subespacios propios de φ relativos a $\lambda_2, \dots, \lambda_p$ ⁽¹⁾. Pero la suma de los subespacios propios de φ es directa, por lo tanto los subespacios propios de φ_1 son necesariamente los subespacios propios de φ relativos a $\lambda_2, \dots, \lambda_p$, y E es la suma directa de los subespacios propios de φ , que son ortogonales dos a dos.

De todo ello resulta que, en una base ortonormal conveniente, la matriz de φ es a la vez diagonal y hermitica, es decir, diagonal con elementos reales. Puesto que los elementos diagonales son valores propios de φ , éstos son reales. c.q.d.

Nota. Se puede ver directamente que los valores propios de φ son reales, y que los subespacios son ortogonales dos a dos. Sea x un vector propio de φ , relativo al valor propio λ . En virtud de (1), se tiene:

$$(\varphi(x) | x) = (\lambda x | x) = (x | \varphi(x)) = (x | \lambda x),$$

o sea

$$\lambda(x | x) = \bar{\lambda}(x | x)$$

de donde

$$\lambda = \bar{\lambda} \text{ ya que } (x | x) > 0.$$

Luego $\lambda \in \mathbb{R}$.

Consideremos ahora dos vectores propios x_1 y x_2 relativos, respectivamente, a los valores propios λ_1, λ_2 , en donde $\lambda_1 \neq \lambda_2$.

⁽¹⁾ En el caso considerado, el hecho de que φ sea diagonalizable implica $p > 1$.

Se tiene:

$$(\varphi(x_1) | x_2) = \lambda_1(x_1 | x_2) = (x_1 | \varphi(x_2)) = \bar{\lambda}_2(x_1 | x_2)$$

de donde (puesto que $\lambda_2 = \bar{\lambda}_2$):

$$(\lambda_1 - \lambda_2)(x_1 | x_2) = 0, \text{ o sea } (x_1 | x_2) = 0, \text{ ya que } \lambda_1 - \lambda_2 \neq 0.$$

Así hemos demostrado que x_1 y x_2 son ortogonales.

Es posible enunciar XIII.6.1 en una forma puramente matricial, que nos será útil en lo sucesivo:

XIII.6.2 *Toda matriz hermítica es diagonalizable, y sus valores propios son reales.*

En particular, toda matriz simétrica **con elementos reales** es diagonalizable (en \mathbf{C}^n) y sus valores propios son reales.

Operadores normales

DEFINICIÓN XIII.6.3

A un endomorfismo $\varphi \in \mathcal{L}(E)$ se le llama **normal** si conmuta con su adjunto, es decir, si verifica:

$$(4) \quad \varphi \circ \varphi^* = \varphi^* \circ \varphi.$$

Si designamos por M la matriz de φ en una base ortonormal, la relación (4) se escribe:

$$(5) \quad M {}^t\bar{M} = {}^t\bar{M} M.$$

Ejemplos

- 1) Si $\varphi = \varphi^*$, (4) se verifica. Un operador autoadjunto es, pues, normal.
- 2) Sea $\varphi \in \mathcal{L}(E)$. La relación

$$\forall x \in E, \forall y \in E, (\varphi(x) | \varphi(y)) = (x | (\varphi^* \circ \varphi)(y))$$

prueba que φ es unitario si, y sólo si, φ es invertible y $\varphi^{-1} = \varphi^*$.

En particular, un operador unitario es normal.

- 3) Sin embargo, existen operadores normales que no son ni autoadjuntos ni unitarios, por ejemplo, el operador de matriz A en \mathbf{C}^2 :

$$A = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}.$$

TEOREMA XIII.6.3

|| (Teorema espectral para operadores normales.)
 || Si $\varphi \in \mathcal{L}(E)$ es un operador normal; φ es diagonalizable, y sus subespacios
 || propios son ortogonales dos a dos. O también, existe una base ortonormal
 || de E , formada por vectores propios de φ .

Antes de abordar la demostración, señalemos por de pronto que (contrariamente a la conclusión de XIII.6.1) los valores propios de φ no son necesariamente reales. Por ejemplo, los valores propios de $A = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}$ son $\lambda = 0$ y $\lambda = 2i$.

Demostración. Por recurrencia sobre $n = \dim(E)$. El teorema es evidente para $n = 1$. Sean $\lambda_1, \dots, \lambda_p$ los valores propios distintos de φ , y E_1, \dots, E_p los subespacios propios correspondientes. Si $E_1 = E$, el teorema está demostrado. Si no, pongamos: $F_1 = E_1^\perp$. Se tiene: $1 \leq \dim(F_1) \leq n - 1$.

Probemos que F_1 es estable para φ . En primer lugar para todo $x \in E_1$, se tiene:

$$\varphi^*(x) \in E_1, \text{ pues } \varphi[\varphi^*(x)] = (\varphi \circ \varphi^*)(x) = (\varphi^* \circ \varphi)(x) = \lambda_1 \varphi^*(x).$$

Si $y \in F_1$, se tiene entonces $(\varphi(y) | x) = (y | \varphi^*(x)) = 0$, ya que $\varphi^*(x) \in E_1$. De donde

$$\varphi(F_1) \subset F_1.$$

Se ha visto anteriormente que F_1 es φ -estable. Ahora veremos que F_1 es φ^* -estable. En efecto, para todo $y \in E_1$ y todo $x \in F_1$, se tiene:

$$(\varphi^*(x) | y) = (x | \varphi(y)) = 0$$

(ya que E_1 es φ -estable), por lo tanto $\varphi^*(x) \in E_1^\perp = F_1$, tal como habíamos enunciado.

La relación (1) muestra entonces que la restricción de φ^* a F_1 es, precisamente, el adjunto de la restricción; por consiguiente, la restricción φ_1 de φ a F_1 es un operador normal de F_1 . En virtud de la hipótesis de recurrencia, F_1 es la suma directa de los subespacios de φ_1 , y estos subespacios son ortogonales dos a dos. Estos subespacios están contenidos a priori en los subespacios propios de φ relativos a $\lambda_2, \dots, \lambda_p$ ⁽¹⁾. Puesto que la suma de los subespacios propios de φ es directa, los subespacios propios de φ_1 son necesariamente E_2, E_3, \dots, E_p , y E es la suma di-

⁽¹⁾ En el caso que se considera, el hecho de que φ_1 sea diagonalizable asegura que $p > 1$.

recta de E_1, E_2, \dots, E_p . Entonces es suficiente tomar una base ortonormal en cada uno de los subespacios E_k ($1 \leq k \leq p$). c.q.d.

Aplicación

Toda matriz unitaria $U \in U(n, \mathbf{C})$ admite una base ortonormal de vectores propios. En dicha base, la matriz V de la transformación representada por U en la base inicial es de la forma:

$$V = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

Pero V debe ser unitaria (por lo tanto normal), pues se verifica: $V = P^{-1}UP$, en donde P es unitaria. Si escribimos ${}^tV \cdot V = I_n$, vemos que los λ_i tienen módulo 1 (lo cual se puede establecer directamente aplicando la relación ${}^tU\bar{U} = I_n$.) Podemos, pues, enunciar:

XIII.6.4 Si u es un automorfismo unitario del espacio prehilbertiano E , existe una base ortonormal de E en que la matriz de u es de la forma

$$\begin{bmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & e^{i\theta_n} \end{bmatrix}.$$

En el § 10 encontraremos aplicaciones importantes de XIII.6.4. En general, el teorema XIII.6.3 permite establecer:

XIII.6.5 Sea $\varphi \in \mathcal{L}(E)$ un operador **normal**. Para que φ sea **autoadjunto** es necesario y suficiente que sus valores propios sean **reales**.
Para que φ sea unitario es necesario y suficiente que sus valores propios tengan módulo 1.

Demostración. Sea (e_1, \dots, e_n) una base ortonormal de E en la que la matriz D de φ sea diagonal.

Entonces, φ es autoadjunto (resp. unitario) si, y sólo si, $D = \bar{D}$ (resp. $D\bar{D} = I_n$). De ahí el resultado. c.q.d.

§ XIII.7 TEORÍA ESPECTRAL (FORMAS CUADRÁTICAS REALES)

● En lo sucesivo, E_n designará un espacio euclídeo de dimensión n . Recordemos (cf. XII.3.2) que, para toda forma lineal u definida en E_n , existe un elemento $y \in E_n$ único tal que, para todo $x \in E_n$, se verifica:

$$u(x) = x \cdot y.$$

Como aplicación, sea $\varphi \in \mathcal{L}(E_n)$ un endomorfismo de E_n , para todo $y \in E_n$, la aplicación $x \mapsto \varphi(x) \cdot y$ es una forma lineal definida en E_n . Existe, pues, un $y' \in E_n$ único tal que, para todo $x \in E_n$, se verifica:

$$\varphi(x) \cdot y = x \cdot y'.$$

DEFINICIÓN XIII.7.1

Para todo endomorfismo $\varphi \in \mathcal{L}(E_n)$, se llama **endomorfismo traspuesto** (o **adjunto**) de φ , y se designa por ${}^t\varphi$ (o φ^*), al endomorfismo definido por

$$(1) \quad \forall x \in E_n, \quad \forall y \in E_n, \quad \varphi(x) \cdot y = x \cdot {}^t\varphi(y).$$

Si, en virtud del teorema XII.3.2 se identifica E_n con su dual E_n^* , el endomorfismo ${}^t\varphi$ se identifica con la *traspuesta* de φ en el sentido de la definición VIII.5.2, y esto justifica la terminología empleada.

Si M y N designan las matrices de φ y de ${}^t\varphi$ en una base ortonormal cualquiera, vemos, como en el § 6, que

$$(2) \quad N = {}^tM.$$

Se comprueba que la aplicación $\varphi \mapsto {}^t\varphi$ es un automorfismo involutivo del espacio vectorial $\mathcal{L}(E_n)$, que verifica ${}^t(\varphi \circ \psi) = {}^t\psi \circ {}^t\varphi$.

DEFINICIÓN XIII.7.2

El endomorfismo $\varphi \in \mathcal{L}(E_n)$ es **simétrico** si $\varphi = {}^t\varphi$, es decir si

$$(3) \quad \varphi(x) \cdot y = x \cdot \varphi(y) \quad \text{para todo } x \in E_n \text{ y todo } y \in E_n.$$

Según (2), un endomorfismo simétrico está caracterizado por el hecho de que su matriz en toda base ortonormal es una *matriz simétrica*.

TEOREMA XIII.7.1

Un operador simétrico $\varphi \in \mathcal{L}(E_n)$ tiene todos sus valores propios reales, es diagonalizable, y sus subespacios propios son ortogonales dos a dos.

|| Con otras palabras, E_n es la suma directa de los subespacios propios de φ , y éstos son ortogonales dos a dos.

Demostración. Elijamos una base ortonormal cualquiera (e_1, \dots, e_n) de E_n . Con la ayuda de esta base, E_n se identifica con \mathbf{R}^n . Por medio de la inyección canónica $\mathbf{R} \rightarrow \mathbf{C}$ (que envía el número real x al número complejo $z = x$), definimos una inyección $\mathbf{R}^n \rightarrow \mathbf{C}^n$, que permite identificar \mathbf{R}^n con un subconjunto de \mathbf{C}^n . Entonces (e_i) es una base de \mathbf{C}^n , y dotamos a \mathbf{C}^n de la estructura prehilbertiana canónica asociada a esta base. La restricción a $\mathbf{R}^n \times \mathbf{R}^n$ del producto escalar hermitico de \mathbf{C}^n no es otro que el producto escalar de E_n .

La matriz M de φ en (e_i) es simétrica y real. Define un operador hermitico $\hat{\varphi}$ de \mathbf{C}^n , cuya restricción a \mathbf{R}^n es φ .

Aplicamos XIII.6.1: los valores propios de $\hat{\varphi}$ son reales, y $\hat{\varphi}$ es diagonalizable. Luego, para todo valor propio λ de M , de multiplicidad α , el rango de la matriz $M - \lambda I_n$ (considerada como matriz de $M_n(\mathbf{C})$) es igual a $n - \alpha$. Como $M - \lambda I_n$ está formado por elementos reales, este rango es también el rango de $M - \lambda I_n$ considerada como matriz de $M_n(\mathbf{R})$ (ver caracterización de rango mediante subdeterminantes, Cap. X). Resulta de todo ello que el subespacio propio de φ , asociado a λ , es de dimensión α (cf. XI.2.3). Por lo tanto, φ es diagonalizable. En virtud de las observaciones del principio, si $x \in E_n$ e $y \in E_n$, las relaciones $x \cdot y = 0$ y $(x | y) = 0$ son equivalentes. De XIII.6.1 resulta que los subespacios propios de φ son ortogonales dos a dos. c.q.d.

Operador asociado a una forma cuadrática

Sea Q una forma cuadrática definida en E_n , y B su forma polar. Designemos por $I: E_n \rightarrow E_n^*$ al isomorfismo canónico (definido por: $(\forall x \in E_n, \forall y \in E_n)$, $\langle x, I(y) \rangle = x \cdot y$) y por $J: E_n \rightarrow E_n^*$ a la aplicación lineal asociada a B , definida por: $(\forall x \in E_n, \forall y \in E_n)$ $\langle x, J(y) \rangle = B(x, y)$.

Se da la siguiente

DEFINICIÓN XIII.7.3

{ Con las notaciones anteriores, al endomorfismo

$$u_Q = I^{-1} \circ J \in \mathcal{L}(E_n)$$

 se le llama **operador asociado** a Q .

Las propiedades siguientes son inmediatas:

1) u_Q es el único endomorfismo de E_n que verifica

$$(\forall x \in E_n, \forall y \in E_n) x \cdot u_Q(y) = B(x, y)$$

2) El endomorfismo u_Q es *simétrico*. En efecto, para todo par $x, y \in E_n$, se tiene, según lo que antecede:

$$x \cdot u_Q(y) = u_Q(x) \cdot y.$$

3) La aplicación $Q \mapsto u_Q$ es un **isomorfismo** del espacio vectorial de las *formas cuadráticas definidas en E_n* , en el espacio vectorial de los *endomorfismos simétricos de E_n* .

4) Si $\mathcal{B} = (e_1, \dots, e_n)$ es una **base ortonormal** de E_n , la matriz de u_Q en \mathcal{B} es igual a la matriz $[B(e_i, e_j)]_{1 \leq i \leq n, 1 \leq j \leq n}$ de Q en \mathcal{B} .

DEFINICIÓN XIII.7.4

$\left\{ \begin{array}{l} \text{Se llama } \mathbf{vector\ principal} \text{ (resp. } \mathbf{dirección\ principal, subespacio prin-} \\ \mathbf{cipal)} \text{ de una forma cuadrática } Q \text{ definida en } E_n, \text{ a todo } \mathbf{vector\ propio} \\ \text{(resp. } \mathbf{dirección\ propia, subespacio propio)} \text{ del operador asociado a } Q. \end{array} \right.$

Aplicación al estudio de las cuádricas

Sea Q una forma cuadrática definida en E_n , y sea $b \in \mathbf{R}$. Se llama «cuádrica de ecuación $Q(x) = b$ » a la terna (Q, b, \mathcal{S}) , en donde \mathcal{S} es el conjunto de los $x \in E_n$ tales que $Q(x) = b$. (\mathcal{S} es, en general, una hipersuperficie, y \mathcal{S} es el *conjunto de los puntos de la cuádrica*.)

A los subespacios principales de Q se les llama los **subespacios principales** de la cuádrica (dependen únicamente de Q , y no de b).

Según XIII.7.1, *existe por lo menos una base ortonormal $\mathcal{B} = (e_1, \dots, e_n$ de E_n formada por los vectores principales de la cuádrica.*

En dicha base \mathcal{B} , si $x = \sum_{i=1}^n x_i e_i$ se tiene,

$$(4) \quad Q(x) = \sum_{i=1}^n \lambda_i x_i^2$$

en donde $(\lambda_1, \dots, \lambda_n)$ son los valores propios de u_Q .

La ecuación $Q(x) = b$ equivale entonces a

$$(5) \quad \sum_{i=1}^n \lambda_i x_i^2 = b.$$

Una relación de la forma (5) se llama **ecuación reducida** de la cuádrica (en una referencia ortonormal y principal). La relación (5) muestra que para todo i ($1 \leq i \leq n$), el hiperplano H_i , de ecuación $x_i = 0$, en \mathcal{B} , es un *hiperplano de simetría de \mathcal{S}* .

Con mayor precisión, sean (ρ_1, \dots, ρ_p) los valores propios distintos de u_Q , y F_1, \dots, F_p los subespacios propios correspondientes. Sea Ω el subgrupo del grupo ortogonal $O(E_n)$ formado por los $u \in O(E_n)$ tales que, para todo i , F_i es u -estable y que el endomorfismo u_i de F_i inducido por u pertenezca a $O(E_i)$. Entonces (5) muestra fácilmente que el conjunto \mathcal{S} es *globalmente invariante para todo $u \in \Omega$* .

Observemos que la aplicación $u \mapsto (u_i)_{1 \leq i \leq p}$ es un isomorfismo de Ω en el grupo producto $O(F_1) \times O(F_2) \times \dots \times O(F_p)$.

Longitud de los ejes de la cuádrica de ecuación $Q(x) = b$

Consideremos de nuevo las notaciones del teorema XIII.7.2. La búsqueda de los valores propios de φ equivale a la de los $\lambda \in \mathbf{R}$ respecto de los que el sistema de ecuaciones lineales y homogéneas

$$(6) \quad \frac{1}{2} \frac{\partial Q}{\partial x_i} = \lambda x_i \quad (1 \leq i \leq n) \quad \text{admite una solución no trivial.}$$

Cuando ocurra, los puntos de \mathcal{S} que verifiquen (6) constituirán los *vértices* de la cuádrica correspondientes al valor propio λ . Si $x = (x_1, \dots, x_n)$ designa uno de dichos puntos, se tiene

$$\frac{1}{2} \sum_{i=1}^n x_i \frac{\partial Q}{\partial x_i} = \lambda \sum_{i=1}^n x_i^2.$$

Pero, en virtud de la identidad de Euler, $\sum_{i=1}^n x_i \frac{\partial Q}{\partial x_i} = 2 Q(x)$, de donde

$$Q(x) = \lambda \sum_{i=1}^n x_i^2.$$

La cantidad $\sum_{i=1}^n x_i^2$ es precisamente el cuadrado L_λ^2 de la longitud del semieje correspondiente a λ . Puesto que $Q(x) = b$, la relación anterior nos proporciona L^2 :

$$(7) \quad \lambda L_\lambda^2 = b.$$

Este resultado también se encuentra evidentemente después de un simple examen de la ecuación reducida de la cuádrica.

Ejemplo 1

En el plano euclídeo \mathbf{R}^2 , hallar la ecuación reducida a sus ejes ortonormales y las direcciones principales de la cónica de ecuación

$$x^2 + 2xy - 3y^2 = 1.$$

La matriz de la forma cuadrática del primer miembro es $M = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix}$.

La ecuación característica $\det(M - \lambda I_2) = 0$ se escribe $\lambda^2 + 2\lambda - 4 = 0$, y las raíces son $-1 \pm \sqrt{5}$.

La ecuación reducida buscada es entonces:

$$(\sqrt{5} - 1) X^2 - (\sqrt{5} + 1) Y^2 = 1.$$

La cónica es una hipérbola, cuyas asíntotas tienen por pendiente

$$\pm \sqrt{\frac{\sqrt{5} - 1}{\sqrt{5} + 1}}$$

respecto de los ejes principales. Estos últimos tienen por dirección las direcciones propias de M , a saber los vectores:

$$f_1 \begin{Bmatrix} 1 \\ -2 - \sqrt{5} \end{Bmatrix}, \quad f_2 \begin{Bmatrix} 1 \\ -2 + \sqrt{5} \end{Bmatrix}.$$

Según (7), el semieje de la hipérbola tiene longitud $L = \frac{1}{\sqrt{\sqrt{5} - 1}}$.

Ejemplo 2

Las mismas cuestiones, en el espacio euclídeo \mathbf{R}^3 , para la cuádrica de ecuación

$$5x^2 + y^2 + z^2 - 2xy + 2xz - 6yz = 1.$$

La matriz de la forma del primer miembro es

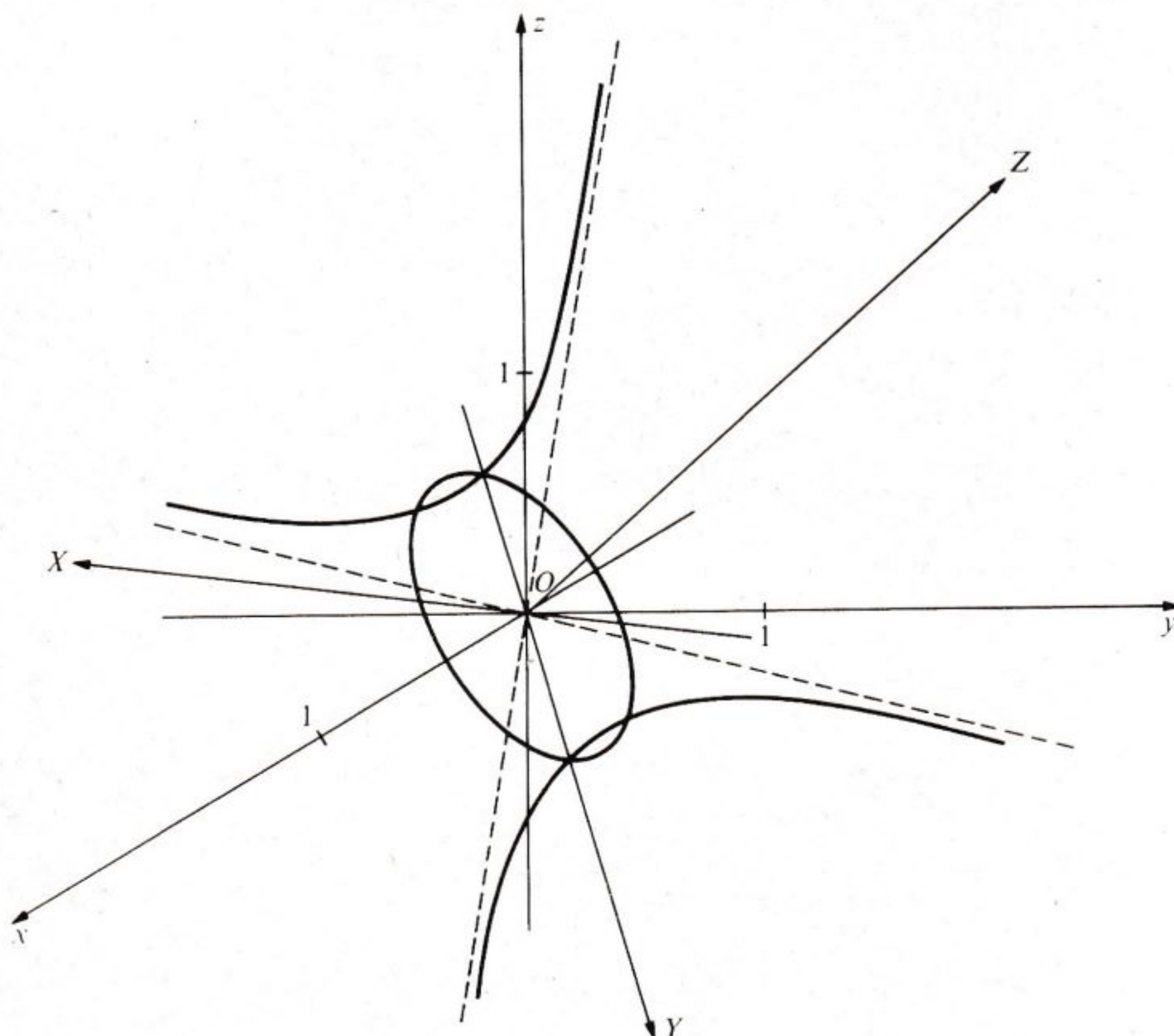
$$M = \begin{bmatrix} 5 & -1 & 1 \\ -1 & 1 & -3 \\ 1 & -3 & 1 \end{bmatrix}, \text{ de ecuación característica } (\lambda - 6)(\lambda^2 - \lambda - 6) = 0, \text{ y va-}$$

lores propios $\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = -2$.

Ecuación reducida: $6X^2 + 3Y^2 - 2Z^2 = 1$. Se trata de un hiperboloide de una hoja, cuyos ejes principales tienen las direcciones de los vectores propios de M , por ejemplo, los vectores:

$$f_1 \begin{Bmatrix} 2 \\ -1 \\ 1 \end{Bmatrix}, \quad f_2 \begin{Bmatrix} 1 \\ 1 \\ -1 \end{Bmatrix}, \quad f_3 \begin{Bmatrix} 0 \\ 1 \\ 1 \end{Bmatrix}.$$

A continuación representamos las secciones del hiperboloide por los planos principales XOY e YOZ .



Según (7), las longitudes de los semiejes son respectivamente $\frac{1}{\sqrt{6}}$, $\frac{1}{\sqrt{3}}$.

§ XIII.8 TEORÍA ESPECTRAL (FORMAS CUADRÁTICAS SOBRE UN CUERPO CUALQUIERA)

Es posible enunciar el teorema espectral XIII.7.1 en la forma siguiente:

Dado, en el \mathbf{R} -espacio vectorial E de dimensión finita n , una forma cuadrática Q_2 definida positiva, y una forma cuadrática Q_1 cualquiera, existe entonces una base de E ortonormal para Q_2 y ortogonal para Q_1 .

En efecto, dotemos a E de la estructura euclídea definida por Q_2 . Sea (e_i) una base ortonormal cualquiera fija de E , y sea φ_1 el operador simétrico asociado a Q_1 en el espacio euclídeo (E, Q_2) . Según XIII.7.1, existe una base ortonormal (ε_i) de E

en la que la matriz de φ_1 es diagonal. Esta base (ε_i) es entonces a la vez ortonormal para Q_2 y ortogonal para Q_1 ; de ahí nuestra afirmación.

De esta manera hemos sido conducidos a un problema general, ligeramente diferente de los problemas tratados en los §§ 6 y 7, que es el de la *reducción simultánea* de dos formas cuadráticas.

En lo sucesivo, K designará un cuerpo conmutativo de característica $\neq 2$, E un K -espacio vectorial de dimensión n , y Q_1, Q_2 dos formas cuadráticas definidas en E . Supondremos que Q_2 es *no degenerada*, y daremos condiciones necesarias y suficientes para que exista una base de E , ortogonal simultáneamente para Q_1 y para Q_2 .

DEFINICIÓN XIII.8.1

Sean Q_1 y Q_2 dos formas cuadráticas, de formas polares B_1 y B_2 , definidas en el espacio vectorial E de dimensión $n \geq 1$, con Q_2 **no degenerada**. Para $i = 1, 2$, sea $J_i : E \rightarrow E^*$ la aplicación lineal tal que: $(\forall x \in E)$, $J_i(x) = B_i(x, \cdot)$. Se llama **operador asociado al par** (Q_1, Q_2) al endomorfismo $\varphi \in \mathcal{L}(E)$ definido por

$$\varphi = J_2^{-1} \circ J_1.$$

El polinomio característico de φ se llama **polinomio de los invariantes** (o por abuso de lenguaje, *ecuación de los invariantes*) del par (Q_1, Q_2) . A sus coeficientes se les llama **invariantes simultáneos** del par (Q_1, Q_2) .

Es inmediato que, si \mathcal{B} es una base cualquiera de E , y si $M_i (i = 1, 2)$ designa la matriz de Q_i en esta base, entonces la matriz de φ en \mathcal{B} es

$$M_2^{-1} \times M_1.$$

El teorema que resuelve nuestro problema es el siguiente:

TEOREMA XIII.8.1

Con las notaciones de la definición XIII.8.1, para que exista una base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E , ortogonal a la vez respecto de Q_1 y de Q_2 , es necesario y suficiente que φ sea diagonalizable.

Demostración

a) Supongamos que existe \mathcal{B} . Veamos entonces que, para todo i ($1 \leq i \leq n$), e_i es un vector propio de φ , lo que demostrará que φ es diagonalizable. Hagamos: $\varepsilon_i = \varphi(e_i) (= J_2^{-1} \circ J_1(e_i))$. Para todo j ($1 \leq j \leq n$), se tiene:

$$B_2(\varepsilon_i, e_j) = \langle J_2(\varepsilon_i), e_j \rangle = \langle J_1(e_i), e_j \rangle = B_1(e_i, e_j).$$

Si $j \neq i$, se deduce: $B_2(\varepsilon_i, e_j) = 0$. Luego, relativamente a B_2 , ε_i es ortogonal a todos los $e_j (j \neq i)$, y puesto que Q_2 es no degenerada, esto implica que ε_i es colineal a e_i , que es lo que queríamos establecer.

b) Recíprocamente, supongamos que φ es diagonalizable. Sean $\lambda_1, \dots, \lambda_p$ sus distintos valores propios, y E_1, E_2, \dots, E_p los subespacios correspondientes. En E_i , sea \mathcal{B}_i una base ortogonal para la restricción de Q_2 a E_i (cf. XII.3.1), y sea \mathcal{B} la base de E obtenida «yuxtaponiendo» los \mathcal{B}_i . Vamos a calcular los $B_1(e_j, e_k)$ y los $B_2(e_j, e_k)$ para $1 \leq j \leq n$, $1 \leq k \leq n$, $j \neq k$ (en donde $\mathcal{B} = (e_1, e_2, \dots, e_n)$).

Primer caso: e_j y e_k pertenecen a un mismo E_i .

Entonces

$$B_2(e_j, e_k) = 0,$$

pues:

$$\begin{aligned} B_1(e_j, e_k) &= \langle J_1(e_j), e_k \rangle = \langle J_2(\varphi(e_j)), e_k \rangle = \\ &= \langle \lambda_i J_2(e_j), e_k \rangle = \lambda_i B_2(e_j, e_k) = 0. \end{aligned}$$

Segundo caso: se tiene: $e_j \in E_{i_1}$ y $e_k \in E_{i_2}$, con $i_1 \neq i_2$.

$$\begin{aligned} \text{Entonces, } B_1(e_j, e_k) &= \langle J_1(e_j), e_k \rangle = \langle J_2(\varphi(e_j)), e_k \rangle = \langle \lambda_{i_1} J_2(e_j), e_k \rangle = \\ &= \lambda_{i_1} B_2(e_j, e_k). \end{aligned}$$

Análogamente,

$$B_1(e_j, e_k) = B_1(e_k, e_j) = \lambda_{i_2} B_2(e_k, e_j) = \lambda_{i_2} B_2(e_j, e_k).$$

De lo que se deduce: $(\lambda_{i_1} - \lambda_{i_2}) B_2(e_j, e_k) = 0$, y $B_2(e_j, e_k) = 0$.

Luego $B_1(e_j, e_k) = \lambda_{i_1} B_2(e_j, e_k) = 0$. Por lo tanto se ha demostrado que \mathcal{B} es a la vez Q_1 -ortogonal y Q_2 -ortogonal.]]

El teorema XIII.8.1 admite un complemento importante:

XIII.8.2. Con las notaciones e hipótesis de la definición XIII.8.1, supongamos que φ es diagonalizable y que Q_1 es no degenerada. Sean $\lambda_1, \dots, \lambda_p$ los valores propios distintos de φ , y E_1, \dots, E_p los subespacios propios correspondientes. Entonces el grupo $O(Q_1) \cap O(Q_2)$ es el conjunto G de los $u \in O(Q_2)$ que dejan estable cada uno de los E_i . Este grupo es canónicamente isomorfo a

$$O_1(Q_2) \times O_2(Q_2) \times \dots \times O_p(Q_2),$$

en donde $O_i(Q_2)$ designa el grupo ortogonal de la restricción de Q_2 a E_i .

Demostración.

a) G es evidentemente un subgrupo de $O(Q_2)$. Demostremos primeramente que:

$$G \subset O(Q_1).$$

Si $u \in G$, calculemos $B_1(u(x), u(x))$ para $x \in E$. Si hacemos $x = \sum_{i=1}^p x_i$ con $x_i \in E_i$ para todo i , se obtiene:

$$\begin{aligned} B_1(u(x), u(x)) &= \langle J_1(u(x)), u(x) \rangle = \langle J_2(\varphi(u(x))), u(x) \rangle \\ &= \sum_{i,j} \lambda_i \langle J_2(u(x_i)), u(x_j) \rangle = \sum_{i,j} \lambda_i B_2(u(x_i), u(x_j)) = \\ &= \sum_{i,j} \lambda_i B_2(x_i, x_j) = \sum_{i,j} \langle J_2 \circ \varphi(x_i), x_j \rangle = B_1(x, x), \end{aligned}$$

lo que demuestra que efectivamente $u \in O(Q_1)$.

b) Recíprocamente, sea $u \in O(Q_1) \cap O(Q_2)$. Demostremos que: $u \in G$. Para ello, establezcamos en primer lugar que $\varphi \circ u = u \circ \varphi$. Para $x \in E$ y todo $y \in E$, se tiene:

$$B_2(\varphi \circ u(x), u(y)) = \langle J_1 \circ u(x), u(y) \rangle = B_1(u(x), u(y)) = B_1(x, y) = \langle J_1(x), y \rangle$$

(puesto que $u \in O(Q_1)$), luego:

$$B_2(u \circ \varphi(x), u(y)) = B_2(\varphi(x), y) = \langle J_1(x), y \rangle, \quad (\text{puesto que } u \in O(Q_2)).$$

Del hecho de no ser Q_2 degenerada, la relación

$$(\forall x \in E), (\forall y \in E) \quad B_2(\varphi \circ u(x), u(y)) = B_2(u \circ \varphi(x), u(y))$$

implica: $\varphi \circ u = u \circ \varphi$.

Una vez establecido esto, si $x \in E_i$, se tiene: $\varphi \circ u(x) = u \circ \varphi(x) = \lambda_i u(x)$, luego

$$u(x) \in \text{Ker}(\varphi - \lambda_i \text{Id}_{E_i}) = E_i.$$

Esto muestra que E_i es u -estable para todo i , de donde: $u \in G$. Hemos demostrado de esta manera que $G = O(Q_1) \cap O(Q_2)$.

c) Finalmente, para todo $u \in G$, sea u_i el elemento de $O_i(Q_2)$ inducido por la restricción de u a E_i . Es inmediato que la aplicación $u \mapsto (u_1, u_2, \dots, u_p)$ es un isomorfismo de G en el grupo producto $O_1(Q_2) \times O_2(Q_2) \times \dots \times O_p(Q_2)$. c.q.d.

Notas

1) Si el cuerpo de base K es algebraicamente cerrado, toda base ortogonal para una forma cuadrática *no degenerada* se puede transformar, por multiplicación de sus elementos por escalares convenientes, en una base ortonormal. Se obtiene entonces, en este caso, un teorema ligeramente más fino que XIII.8.1:

TEOREMA XIII.8.3

Sea φ el operador asociado a un par de formas cuadráticas (Q_1, Q_2) (en donde Q_2 es no degenerada) en el espacio E de dimensión finita n ; el cuerpo base se supone algebraicamente cerrado. Para que exista una base (e_i) de E ortonormal para Q_2 y ortogonal para Q_1 , es necesario y suficiente que φ sea diagonalizable.

En tal base se tiene, entonces:

$$Q_2(\sum x_i e_i) = \sum_{i=1}^n x_i^2,$$

es decir, $M_2 = I_n$. La matriz diagonal M_1 es también la matriz de φ en (e_i) , y se tiene

$$Q_1(\sum x_i e_i) = \sum_{i=1}^n \lambda_i x_i^2.$$

Los λ_i designan los invariantes del par (Q_1, Q_2) , que son también los elementos diagonales de M_1 .

2) Es posible establecer una teoría análoga con formas hermiticas.

3) Apliquemos XIII.8.2 a la situación de XIII.7.1. En el cuerpo \mathbf{R} , sabemos que dos formas cuadráticas con n variables son simultáneamente diagonalizables cuando una de ellas es definida positiva. En forma matricial, XIII.8.2 nos da entonces el siguiente resultado:

Si M es una matriz simétrica definida positiva, y si N es una matriz simétrica, la matriz MN es diagonalizable.

Este resultado es, por otra parte, equivalente a XIII.7.1.

Hay un enunciado análogo para las matrices hermiticas.

4) La ecuación de los invariantes de (Q_1, Q_2) depende únicamente del par (Q_1, Q_2) , y cada uno de sus coeficientes expresa, pues, una propiedad intrínseca de este par. En particular, la *traza* de $M_2^{-1} M_1$ es un invariante del par (Q_1, Q_2) cuya interpretación presenta un gran interés (cf. ejercicios).

Interpretación geométrica cuando $K = \mathbf{C}$

Si Q es una forma cuadrática definida en \mathbf{C}^{n+1} , la ecuación $Q(x) = 0$ define una cuádrica \mathcal{S} del espacio proyectivo $\mathcal{P}_n(\mathbf{C})$ ⁽¹⁾. Dos subespacios vectoriales de \mathbf{C}^{n+1} , V, W de dimensiones $p+1, q+1$, se proyectan sobre $\mathcal{P}_n(\mathbf{C})$ según dos variedades proyectivas de dimensiones p, q , a saber: v, w . Decir que V y W son ortogonales respecto de Q , significa que las variedades v y w son *conjugadas* respecto de la cuádrica \mathcal{S} .

Sean Q_1 y Q_2 dos formas cuadráticas definidas en \mathbf{C}^{n+1} que definan dos cuádricas \mathcal{S}_1 y \mathcal{S}_2 de $\mathcal{P}_n(\mathbf{C})$. Decir que Q_2 es no degenerada significa que \mathcal{S}_2 es *propia*, es decir, *sin puntos singulares* (cf. tomos 2 y 3).

El conjunto de las cuádricas de ecuación $Q_1 - \lambda Q_2 = 0$ ($\lambda \in \mathbf{C}$), aumentado con la cuádrica $Q_2 = 0$, forma el *haz lineal de cuádricas de base* (Q_1, Q_2) .

Establecido esto, una base (e_1, \dots, e_{n+1}) de \mathbf{C}^{n+1} ortogonal a la vez para Q_1 y Q_2 , nos da en $\mathcal{P}_n(\mathbf{C})$ los $n+1$ puntos (a_1, \dots, a_{n+1}) proyectivamente libres, tales que, para todo i , el hiperplano polar de a_i respecto de \mathcal{S}_2 es el engendrado por los

⁽¹⁾ Suponemos conocidas las nociones elementales de geometría proyectiva, cf. tomo 3.

$(a_j)_{j \neq i}$. A tal sistema se le llama *autopolar respecto a \mathcal{S}_2* . Si \mathcal{S}_1 es propio, el sistema es también autopolar respecto a \mathcal{S}_1 . En resumen, la formulación geométrica del problema de la reducción simultánea de Q_1 y Q_2 es la siguiente:

Dadas dos cuádricas propias \mathcal{S}_1 y \mathcal{S}_2 en el espacio proyectivo $\mathcal{P}_n(\mathbf{C})$, buscar los sistemas de $n + 1$ puntos de $\mathcal{P}_n(\mathbf{C})$, autopolares a la vez respecto a \mathcal{S}_1 y a \mathcal{S}_2 .

Nota. Un sistema autopolar respecto a \mathcal{S}_1 y a \mathcal{S}_2 es también autopolar respecto a toda cuádrica del haz de base $\mathcal{S}_1, \mathcal{S}_2$.

Falta interpretar los invariantes del par (Q_1, Q_2) . Estos invariantes son las raíces de la ecuación $\det(M_1 - \lambda M_2) = 0$. Esta ecuación determina las cuádricas $Q_1 - \lambda Q_2 = 0$ que son no propias, y los puntos singulares de estas cuádricas son los puntos de las «rectas propias» del operador. Por una simple transcripción del lenguaje, el teorema XIII.8.2 se puede enunciar de la siguiente manera:

TEOREMA XIII.8.4

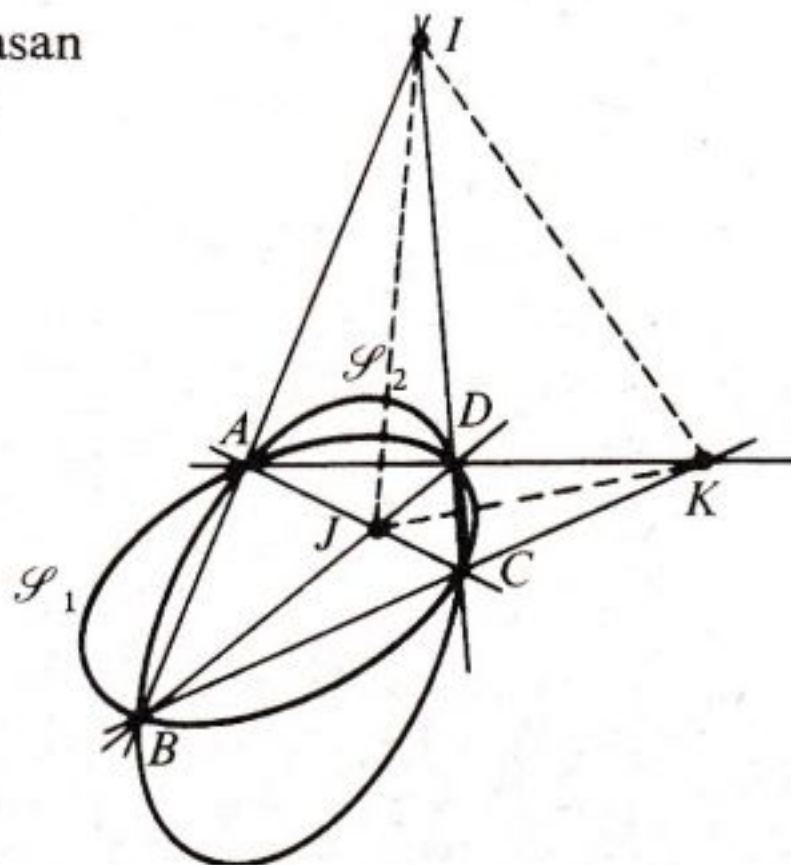
Sean $\mathcal{S}_1, \mathcal{S}_2$ dos cuádricas propias del espacio proyectivo $\mathcal{P}_n(\mathbf{C})$, de ecuaciones $Q_1(x) = 0, Q_2(x) = 0$ en \mathbf{C}^{n+1} .
Para que existan sistemas de $n + 1$ puntos de $\mathcal{P}_n(\mathbf{C})$, autopolares respecto a \mathcal{S}_1 y a \mathcal{S}_2 a la vez, es necesario y suficiente que el operador φ asociado al par (Q_1, Q_2) sea diagonalizable.

Caso particular. Si φ tiene todos sus valores propios distintos, los subespacios propios de φ son de dimensión 1. Las cuádricas degeneradas del haz $(\mathcal{S}_1, \mathcal{S}_2)$ tienen un único punto singular: son conos (no descompuestos si $n \geq 3$). Los vértices de estos conos forman el único sistema autopolar respecto a \mathcal{S}_1 y a \mathcal{S}_2 , a la vez (salvo para una permutación).

Ejemplo: $n = 2$.

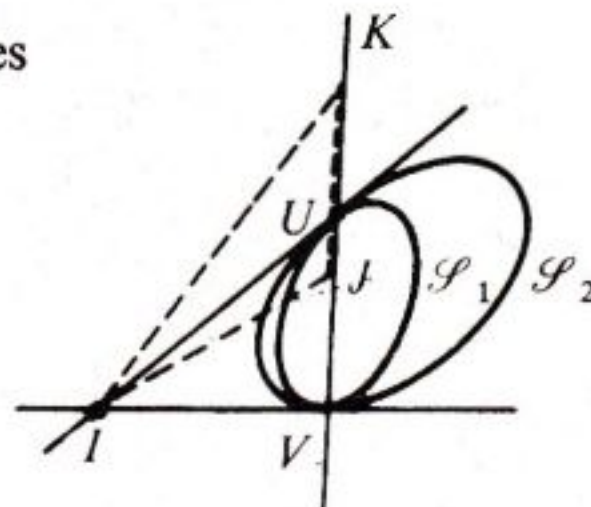
1) \mathcal{S}_1 y \mathcal{S}_2 son cónicas. Si en el haz \mathcal{F} de base \mathcal{S}_1 y \mathcal{S}_2 hay exactamente tres cónicas degeneradas, los tres puntos dobles de ellas constituyen el único triángulo conjugado respecto de \mathcal{S}_1 y de \mathcal{S}_2 , a la vez.

(Cónicas que pasan por A, B, C, D)



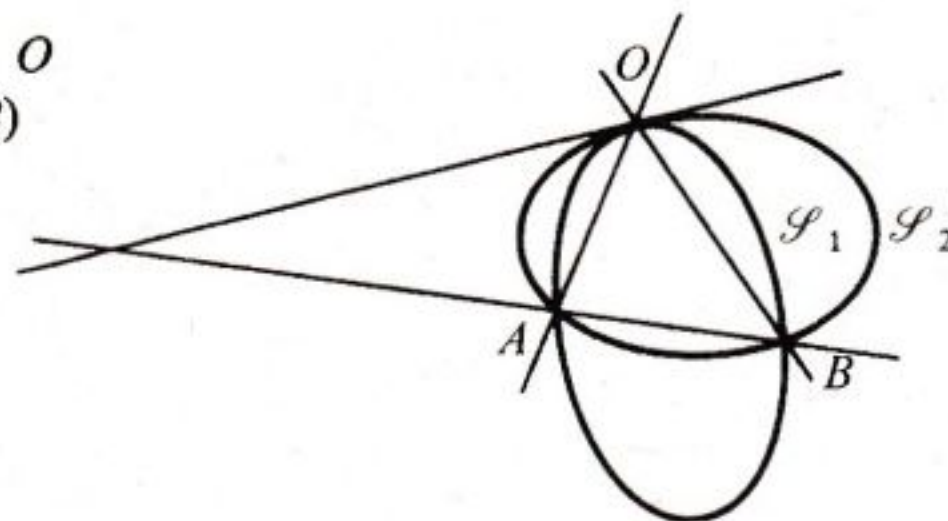
2) Si en \mathcal{F} hay dos cónicas degeneradas, el operador φ es diagonalizable si, y sólo si, \mathcal{F} es un haz de cónicas bitangentes. Existe una infinidad de triángulos autopolares comunes a \mathcal{S}_1 y a \mathcal{S}_2 , tales como IKK (cf. fig.). Basta con tomar J y K sobre la «recta doble» de \mathcal{F} , conjugados respecto de U y V .

(Cónicas bitangentes
en U y V)

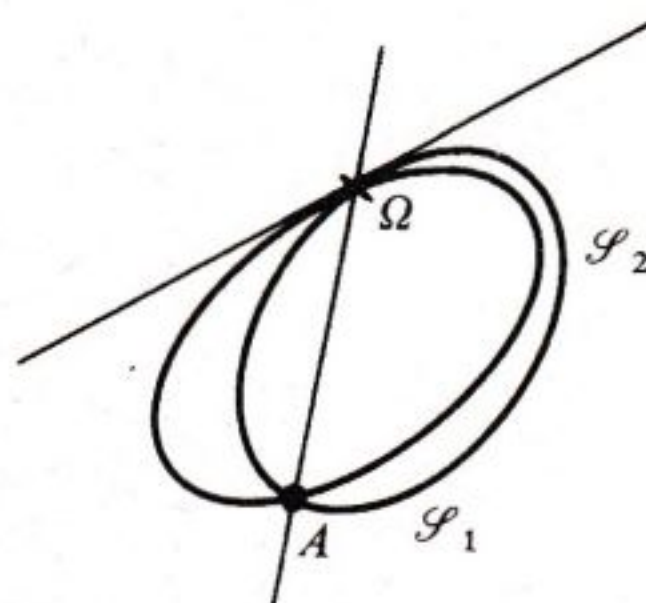


Cuando φ no es diagonalizable, \mathcal{F} es del tipo siguiente (cf. fig.). No existe ningún triángulo conjugado común a \mathcal{S}_1 y a \mathcal{S}_2 .

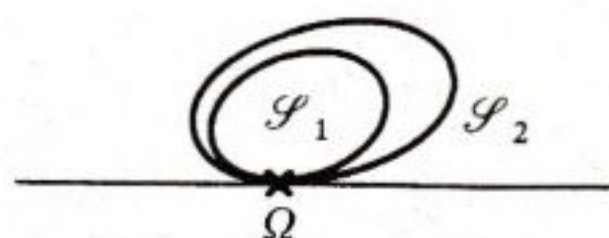
(Cónicas tangentes en O
que pasan por A y B)



3) Si \mathcal{F} posee sólo una cónica degenerada (siendo \mathcal{S}_1 y \mathcal{S}_2 distintos) φ no es diagonalizable; en los dos últimos tipos de haces de cónicas no se dispone de triángulo conjugado común:



(Cónicas osculadoras en Ω que pasan por A)



(Cónicas osculadoras sobre Ω)

§ XIII.9 ISOMETRÍAS DE E_n (ESPACIO EUCLÍDEO DE DIMENSIÓN n)

● En lo sucesivo consideraremos únicamente (salvo mención expresa de lo contrario) espacios euclídeos de dimensión n , por lo tanto isomorfos a E_n (ver Cap. XII, § 5). Los supondremos dotados de la *estructura afín* asociada canónicamente al espacio vectorial subyacente (Cap. VIII, § 6) y se le denominará *espacios afines euclídeos*.

DEFINICIÓN XIII.9.1

Una **isometría** del espacio afín euclídeo E_n es una aplicación $\varphi : E_n \rightarrow E_n$, que conserva las distancias, es decir, tal que, para todo $x \in E_n$ y todo $y \in E_n$, se cumple:

$$(1) \quad |x - y| = |\varphi(x) - \varphi(y)|.$$

Propiedades inmediatas

- Una isometría es inyectiva, pues $\varphi(x) = \varphi(y)$ implica $|\varphi(x) - \varphi(y)| = 0$, de donde $|x - y| = 0$ según (1), es decir $x = y$ ⁽¹⁾.
- La aplicación idéntica es una isometría.
- La compuesta de un número cualquiera de isometrías es una isometría ⁽²⁾.

Ejemplos de isometrías

Las aplicaciones lineales ortogonales de E_n , las traslaciones de E_n . La simetría respecto de un hiperplano de E_n .

El conjunto de los endomorfismos ortogonales de E_n , y el conjunto de las traslaciones de E_n , engendran un subgrupo $\mathcal{O}(E_n)$ del grupo de las biyecciones de E_n . Vamos a demostrar primeramente que *toda isometría de E_n es un elemento de $\mathcal{O}(E_n)$* ; y después estudiaremos más detalladamente este grupo $\mathcal{O}(E_n)$.

Observemos que es posible conocer de forma sencilla todas las isometrías de E_n , si se conocen las que dejan fijo el origen. En efecto si, $\varphi : E_n \rightarrow E_n$ es una isometría, las traslaciones $\tau : y \mapsto \varphi(0) + y$ y $\tau^{-1} : y \mapsto -\varphi(0) + y$ son isometrías; por lo tanto, la aplicación $\psi = \tau^{-1} \circ \varphi$ es una isometría, y se tiene: $\psi(0) = 0$, $\varphi = \tau \circ \psi$. La descomposición de φ en esta forma es además única.

De esta manera hemos sido conducidos a estudiar las isometrías que dejan fijo el origen.

⁽¹⁾ Se observará que no es evidente, a priori, que una isometría sea epiyectiva (es decir, biyectiva) y además es falso en un espacio prehilbertiano real y de dimensión infinita.

⁽²⁾ Las isometrías de un espacio afín euclídeo se estudiarán detalladamente en el tomo 3.

TEOREMA XIII.9.1

Sea φ una isometría del espacio euclídeo afín E_n tal que $\varphi(0) = 0$. Entonces φ es una aplicación lineal. Con otras palabras, φ es un endomorfismo ortogonal de E_n ; en particular, es una aplicación biyectiva.

Demostración. Vamos a demostrar un resultado más general que XIII.9.1 (pues la demostración no es excesivamente más difícil), a saber:

«Sea Q una forma cuadrática *no degenerada* definida en un espacio vectorial E de dimensión finita, y sea $\varphi : E \rightarrow E$ una aplicación tal que $\varphi(0) = 0$ y $Q(x - y) = Q[\varphi(x) - \varphi(y)]$ para todo $x \in E$ y todo $y \in E$ («isometría» respecto de Q). Entonces φ es una biyección lineal de E ».

A este fin, designemos por $B(x, y)$ la forma polar de Q .

a) En primer lugar demostraremos que φ conserva el producto escalar, es decir, que para todo $x \in E$ y todo $y \in E$, se tiene:

$$(2) \quad B(\varphi(x), \varphi(y)) = B(x, y) .$$

Para ello utilizaremos la igualdad: $Q(x - y) = Q(\varphi(x) - \varphi(y))$, que, desarrollada, se escribe:

$$(3) \quad Q(x) + Q(y) - 2 B(x, y) = Q(\varphi(x)) + Q(\varphi(y)) - 2 B(\varphi(x), \varphi(y)) .$$

Puesto que $\varphi(0) = 0$, se tiene además:

$$(4) \quad Q(x) = Q(\varphi(x)) , \quad Q(y) = Q(\varphi(y)) .$$

Por lo tanto de (3) deducimos (2).

b) Sea (e_1, \dots, e_n) una base ortogonal de E . Si hacemos $a_i = Q(e_i)$, obtendremos al $a_i \neq 0$ para $1 \leq i \leq n$, puesto que Q es no degenerada. Tenemos

$$(5) \quad B(\varphi(e_i), \varphi(e_j)) = B(e_i, e_j) \quad (1 \leq i, j \leq n) .$$

Resulta de esto que los vectores $\varphi(e_i)$ ($i = 1, 2, \dots, n$) son ortogonales dos a dos; y para demostrar que $(\varphi(e_i))$ es una *base* ortogonal, es suficiente probar que el sistema $(\varphi(e_i))_{1 \leq i \leq n}$ es un sistema libre. Supongamos entonces que existe una rela-

ción $\sum_{j=1}^n \lambda_j \varphi(e_j) = 0$, que para $1 \leq i \leq n$, implica

$$B \left[\left(\sum_{j=1}^n \lambda_j \varphi(e_j) \right), \varphi(e_i) \right] = 0 = \sum_{j=1}^n \lambda_j B[\varphi(e_j), \varphi(e_i)] = \sum_{j=1}^n \lambda_j B[e_j, e_i] = a_i \lambda_i ,$$

de donde $\lambda_i = 0$ puesto que $a_i \neq 0$.

c) Veamos, finalmente, que φ es lineal. Para todo $x \in E$, se tiene:

$$x = \sum_{i=1}^n x_i e_i, \quad \text{con} \quad x_i = \frac{1}{a_i} B(x, e_i).$$

Asimismo, como $(\varphi(e_i))$ es una base ortogonal y $Q(\varphi(e_i)) = a_i$, se tiene:

$$\varphi(x) = \sum_{i=1}^n y_i \varphi(e_i), \quad \text{con} \quad y_i = \frac{1}{a_i} B(\varphi(x), \varphi(e_i)).$$

En virtud de (5) vemos que $y_i = x_i$, lo cual demuestra la linealidad de φ . Además, puesto que φ transforma (e_i) en una base, φ es biyectiva. c.q.d.

Se observará que el razonamiento anterior es válido para un cuerpo base conmutativo, de característica $\neq 2$, cualquiera.

La generalización de XIII.9.1 así obtenida es interesante, incluso desde el punto de vista de la Física. Por ejemplo, en el espacio vectorial \mathbb{R}^4 , consideramos la forma cuadrática $x^2 + y^2 + z^2 - t^2$; toda «isometría» relativa a esta forma es necesariamente una aplicación lineal. Por lo tanto, las transformaciones físicas que conservan la «función relativista»

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - (t_1 - t_2)^2$$

son biyecciones lineales, es decir, transformaciones de Lorentz. La linealidad de las fórmulas de transformación de coordenadas, cuando se pasa de un sistema de Galileo a otro, *no es un axioma*, sino que es una consecuencia del hecho de que estas fórmulas tengan que conservar la distancia relativista de dos puntos cualesquiera.

DEFINICIÓN XIII.9.2

$\left\{ \begin{array}{l} \text{A las isometrías generales del espacio euclídeo } E_n \text{ se les llama } \mathbf{isome-} \\ \mathbf{trías afines, o isometrías. A las isometrías de } E_n \text{ que dejan fijo} \\ \text{el origen se les llama } \mathbf{isometrías lineales, o isometrías vectoriales.} \end{array} \right.$

En virtud de XIII.9.1, las isometrías lineales de E_n son los elementos del grupo ortogonal $O(E_n)$; y las isometrías afines de E_n forman el subgrupo $\mathcal{O}(E_n)$ del grupo afín de E_n , engendrado por el grupo ortogonal de E_n y el grupo de las traslaciones de E_n .

A los elementos de $SO(E_n)$ (resp. $O(E_n) \setminus SO(E_n)$) también se les llama *isometrías vectoriales directas* (resp. *isometrías vectoriales indirectas*).

Para cerrar este §, señalemos que la teoría anterior se puede extender en parte al caso en que E no es de dimensión finita. Con más precisión, sea E un espacio prehilbertiano real, y designemos siempre por $x.y$ el producto escalar. Vamos a demostrar el siguiente resultado: *toda isometría $\varphi: E \rightarrow E$ tal que $\varphi(0) = 0$, es una aplicación lineal.*

De la misma manera que en el apartado a) de la demostración de XIII.9.1 comprobamos que φ conserva el producto escalar. Para demostrar la linealidad de φ , sean x_1, x_2 dos elementos de E , λ_1, λ_2 dos reales, y consideremos el vector

$$Z = \varphi(\lambda_1 x_1 + \lambda_2 x_2) - \lambda_1 \varphi(x_1) - \lambda_2 \varphi(x_2).$$

Para todo $x \in E$, se tiene:

$$\begin{aligned} Z \cdot \varphi(x) &= \varphi(\lambda_1 x_1 + \lambda_2 x_2) \cdot \varphi(x) - \lambda_1 \varphi(x_1) \cdot \varphi(x) - \lambda_2 \varphi(x_2) \cdot \varphi(x) \\ &= (\lambda_1 x_1 + \lambda_2 x_2) \cdot x - (\lambda_1 x_1) \cdot x - (\lambda_2 x_2) \cdot x = 0. \end{aligned}$$

Por lo tanto, Z es ortogonal a $\varphi(E)$, lo que implica que es ortogonal a $\text{Vect}(\varphi(E))$. En particular $Z \cdot Z = |Z|^2 = 0$, de donde $Z = 0$ (puesto que $Z \in \text{Vect} \varphi(E)$), lo que establece nuestro resultado. c.q.d.

§ XIII.10 ISOMETRIAS VECTORIALES

En el espacio euclídeo E_n , llamaremos *isometría involutiva* a toda isometría vectorial φ , tal que $\varphi^2 = e$ (en donde e designa la aplicación idéntica de E_n). Es fácil conocer la naturaleza de estos automorfismos, cuyos únicos valores propios posibles son $-1, +1$. Llamaremos E_n^-, E_n^+ a los subespacios propios asociados, y escribimos

$$x = \frac{1}{2}(x^+ + x^-), \text{ en donde } x^+ = x + \varphi(x) \text{ y } x^- = x - \varphi(x),$$

Vemos que E_n es la suma directa de E_n^- y E_n^+ , puesto que $x^+ \in E_n^+$ y $x^- \in E_n^-$. Por otra parte, para todo $x \in E_n^+$ y todo $y \in E_n^-$, se tiene:

$$x \cdot y = \varphi(x) \cdot \varphi(y) = \varphi(x) \cdot (-y) = -x \cdot y,$$

de donde, $x \cdot y = 0$, por lo que E_n^+ y E_n^- son suplementarios ortogonales. Con otras palabras:

TEOREMA XIII.10.1

|| Toda isometría involutiva φ de un espacio euclídeo E es una simetría respecto del subespacio E^+ de sus puntos dobles ⁽¹⁾.

⁽¹⁾ Este teorema permanece válido si E es prehilbertiano real de dimensión infinita (demostración idéntica).

Consideremos una base ortonormal (e_1, \dots, e_n) de E , tal que (e_1, \dots, e_p) sea una base ortonormal de E^+ , y (e_{p+1}, \dots, e_n) una base ortonormal de E^- . En esta base, la matriz de φ es:

$$p \left\{ \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ 0 & \dots & 1 & \vdots \\ \vdots & & \ddots & \vdots \\ \vdots & & & -1 \\ 0 & \dots & 0 & -1 \end{bmatrix} \right.$$

Se observa entonces que φ es ortogonal directa si $n - p$ es par, ortogonal indirecta si $n - p$ es impar.

Cuando $p = n - 1$, σ es una *simetría respecto de un hiperplano* o, más brevemente, una *simetría*. Una simetría es siempre ortogonal indirecta. Cuando $p = n - 2$ ($n \geq 2$) a φ se le llama *giro* y entonces es ortogonal directa.

En lo sucesivo, vamos a ver que el conjunto de las simetrías engendra el grupo $O(n)$ de las isometrías vectoriales (para $n \geq 1$), y que, para $n \geq 3$, el conjunto de los giros engendra el grupo $SO(n)$ de las isometrías vectoriales directas.

Para ello establecemos el siguiente teorema de base:

TEOREMA XIII.10.2

Sea φ una isometría lineal de un espacio euclídeo E de dimensión n . Existen subespacios $E_1, \dots, E_q, F_1, \dots, F_r$ de E , ortogonales dos a dos, cuya suma es directa e igual a E y que verifican las siguientes condiciones:

- $\dim E_i = 2, \dim F_j = 1$ ($1 \leq i \leq q, 1 \leq j \leq r$);
- la restricción φ'_i de φ a todo E_i es una rotación plana;
- la restricción φ''_j de φ a todo F_j es una isometría (es decir, la identidad, o una simetría central).

En una base ortonormal de E adaptada a la descomposición de E en la suma directa de los E_i y de los F_j , se introducirá la siguiente notación:

$$M_k = \begin{bmatrix} \cos \theta_k & -\sin \theta_k \\ \sin \theta_k & \cos \theta_k \end{bmatrix} \quad \begin{matrix} \varepsilon_1 = \varepsilon_2 = \dots = \varepsilon_{r_1} = -1, \\ \varepsilon_{r_1+1} = \dots = \varepsilon_r = +1. \end{matrix}$$

Entonces la matriz M de φ tiene la forma:

$$(1) \quad M = \begin{bmatrix} & & 0 & 0 & \dots & \dots & \dots & 0 \\ & M_1 & & & & & & \\ & & 0 & 0 & & & & \\ & 0 & 0 & & & & & \\ & 0 & 0 & & & & & \\ & \vdots & & & & & & \\ & & & & M_q & & & \\ & & & & & 0 & & \\ & & & & & 0 & & \\ & & & & & & \varepsilon_1 & \\ & & & & & & & \varepsilon_{r_1} \\ & & & & & & & \varepsilon_{r_1+1} \\ & & & & & & & 0 \\ & & & & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 & \dots & 0 & \varepsilon_r \end{bmatrix}$$

Demostración del teorema XIII.10.2. Sea (e_1, \dots, e_n) una base ortonormal de E . Con la ayuda de los (e_i) , E se identifica con \mathbf{R}^n . Consideramos \mathbf{R}^n como un \mathbf{R} -subespacio vectorial de \mathbf{C}^n ⁽¹⁾ y dotamos a \mathbf{C}^n de su estructura hermitica canónica. La restricción a $\mathbf{R}^n \times \mathbf{R}^n$ del producto escalar $(x|y)$ es el producto escalar euclídeo de E . Excepcionalmente, lo designaremos por $(x|y)$.

La matriz M de φ en (e_1, \dots, e_n) es ortogonal real. La hacemos actuar sobre \mathbf{C}^n , y define allí un operador unitario Φ . Los valores propios no reales de Φ son conjugados dos a dos, los valores propios reales son iguales a ± 1 . Puesto que todos estos valores propios tienen módulo 1 (cf. T. XIII.6.4), podemos colocarlos en el orden que sigue:

$$(2) \quad (e^{i\theta_1}, e^{-i\theta_1}, e^{i\theta_2}, e^{-i\theta_2}, \dots, e^{i\theta_q}, e^{-i\theta_q}, \varepsilon_1, \dots, \varepsilon_{r_1}, \varepsilon_{r_1+1}, \dots, \varepsilon_r)$$

$$\varepsilon_1 = \dots = \varepsilon_{r_1} = -1,$$

$$\varepsilon_{r_1+1} = \dots = \varepsilon_r = +1.$$

Según el teorema XIII.6.3, existe una base ortonormal de vectores propios de Φ . Los subespacios propios correspondientes a dos valores propios conjugados no reales son ortogonales; es posible, por lo tanto, elegir vectores propios conjugados dos a dos, que constituyan una base de estos subespacios. Finalmente, los vectores propios correspondientes a los ε_i se pueden tomar reales. De esta manera se obtiene una base ortonormal de \mathbf{C}^n :

$$f_1, \bar{f}_1; f_2, \bar{f}_2; \dots; f_q, \bar{f}_q; \quad g_{2q+1}, g_{2q+2}, \dots, g_n,$$

en donde f_k, \bar{f}_k y g_l son vectores propios relativos, respectivamente, a $e^{i\theta_k}, e^{-i\theta_k}; \varepsilon_l$, siendo el vector \bar{f}_k conjugado de f_k , y g_l real ($k \leq q, l > 2q$).

Definamos por medio de:

$$(3) \quad g_{2k-1} = \frac{1}{\sqrt{2}} (f_k + \bar{f}_k), \quad g_{2k} = \frac{i}{\sqrt{2}} (f_k - \bar{f}_k) \quad (1 \leq k \leq q);$$

⁽¹⁾ Esta inyección de \mathbf{R}^n en \mathbf{C}^n se precisó ya en el transcurso de la demostración de XIII.7.1.

Entonces $(g_1, \dots, g_{2q}, g_{2q+1}, \dots, g_n)$ es una base ortogonal y real de \mathbf{C}^n . Además, para $1 \leq k \leq q$ se tiene $(f_k | \bar{f}_k) = 0$, de donde:

$$|g_{2k-1}|^2 = \frac{1}{2} |f_k + \bar{f}_k|^2 = 1, \quad |g_{2k}|^2 = \frac{1}{2} |f_k - \bar{f}_k|^2 = 1$$

Luego (g_1, \dots, g_n) es una base ortonormal real de \mathbf{C}^n , es decir, una base ortonormal de E . La matriz de Φ en esta base se deduce de (3) y de las fórmulas inversas:

$$f_k = \frac{1}{\sqrt{2}} (g_{2k-1} - ig_{2k}), \quad \bar{f}_k = \frac{1}{\sqrt{2}} (g_{2k-1} + ig_{2k}).$$

Se escribe exactamente en la forma (1) que habíamos enunciado de antemano.

Pero (g_1, \dots, g_n) es una base de \mathbf{R}^n , por lo tanto, la matriz de Φ en esta base es precisamente la matriz de φ en esta base; de donde resulta nuestro teorema, si tomamos:

el plano engendrado por $\{g_{2k-1}, g_{2k}\}$ como E_k $1 \leq k \leq q$,
y la recta engendada por g_{2q+l} como F_l $1 \leq l \leq r = n - 2q$.]]

Notas

1) Si n es impar y φ es directa, 1 es valor propio de φ . Volvamos en efecto a la matriz M definida por (1). Se tiene:

$$\det(M) = (-1)^{r_1} = 1,$$

luego r_1 es par; puesto que n posee la paridad de r , $r - r_1$ es impar, por lo tanto, $r - r_1 \neq 0$, de donde resulta nuestra afirmación.

2) Cuando $n = 2$, la demostración anterior prueba que, en la base (g_1, g_2) , la matriz de φ tiene una de las formas siguientes

$$M = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad \text{o} \quad M = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \quad \varepsilon_1 = \pm 1, \varepsilon_2 = \pm 1.$$

Con otras palabras, φ es entonces una rotación (de ángulo eventualmente igual a π) o una simetría. Se hallan de nuevo los resultados de la discusión del capítulo XII, § 7.

Generadores de SO(2)

En virtud de la anterior nota 2), se tiene en primer lugar el resultado siguiente:

XIII.10.3 — Toda isometría vectorial indirecta del plano euclídeo es una simetría.

|| — Toda isometría vectorial directa del plano euclídeo es una rotación.

COROLARIO

|| Toda isometría vectorial directa del plano euclídeo es, de una infinidad de maneras, el producto de dos simetrías. Uno de los ejes puede ser elegido de forma arbitraria.

Demostración. Con la ayuda de una base ortonormal del plano euclídeo, identificamos el plano euclídeo con el plano «complejo» de Cauchy (cf. tomo de Análisis); el número complejo $z = x + iy$ corresponde al punto de coordenadas (x, y) .

La rotación φ de ángulo θ alrededor del origen se halla representada por la transformación $z \mapsto Z(z) = e^{i\theta} z$; mientras que la simetría s , cuyo eje tiene a α como ángulo polar, está representada por $z \mapsto s(z) = e^{2i\alpha} \bar{z}$.

Sean θ y α dos reales cualesquiera. Podemos escribir:

$$e^{i\theta} z = e^{2i(\alpha + \theta/2)} \bar{z}_1, \quad \text{con} \quad z_1 = e^{2i\alpha} \bar{z}.$$

Estas fórmulas prueban que la rotación de ángulo θ se escribe:

$$\varphi = t \circ s,$$

en donde s es la simetría respecto del eje de ángulo polar α , y en donde t es la simetría respecto del eje de ángulo polar $\alpha + \theta/2$. c.q.d.

Estudio de $O(3)$ y $SO(3)$

Según XIII.10.2, una isometría lineal φ de E_3 admite por matriz, en una base ortonormal conveniente $(\vec{i}, \vec{j}, \vec{k})$, una de las matrices siguientes:

$$\begin{aligned} \text{si } \varphi \text{ es indirecta } (\varphi \in O(3) \setminus SO(3)) \quad M &= \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{bmatrix}, \\ \text{si } \varphi \text{ es directa } (\varphi \in SO(3)) \quad M &= \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

En el segundo caso, la recta Δ engendrada por \vec{k} es fija en cada punto, y la restricción φ' de φ al plano (\vec{i}, \vec{j}) es una rotación plana. En virtud del corolario de XIII.10.3, φ' se puede descomponer en el producto de dos simetrías s', t' respecto a rectas S, T del plano (\vec{i}, \vec{j}) , de una infinidad de formas diferentes. Desig-

nemos por s y t las simetrías respecto de los planos (S, Δ) y (T, Δ) , y vemos que $\varphi = s \circ t$. Luego φ se puede descomponer (de una infinidad de maneras) en el producto de dos simetrías planas. Pero se tiene también: $\varphi = s_1 \circ t_1$, si designamos por s_1 y t_1 los giros respecto de S y T . Luego φ se puede descomponer también (de una infinidad de maneras) en el producto de dos giros.

En el primer caso ($\varphi \in O(3) \setminus SO(3)$), el plano (\vec{i}, \vec{j}) es globalmente invariante por φ , y la restricción de φ a este plano es una rotación plana. Conservemos las notaciones anteriores, e introduzcamos la simetría u respecto del plano (\vec{i}, \vec{j}) . Vemos que $\varphi = s \circ t \circ u = u \circ s \circ t$. Luego en este caso, φ es el producto de tres simetrías planas.

Rotaciones de \mathbf{R}^4

Siempre en virtud de XIII.10.2, una isometría lineal directa φ de \mathbf{R}^4 admite como matriz, en una base ortonormal conveniente,

$$M = \begin{bmatrix} \cos \theta_1 & -\operatorname{sen} \theta_1 & 0 & 0 \\ \operatorname{sen} \theta_1 & \cos \theta_1 & 0 & 0 \\ 0 & 0 & \cos \theta_2 & -\operatorname{sen} \theta_2 \\ 0 & 0 & \operatorname{sen} \theta_2 & \cos \theta_2 \end{bmatrix}$$

φ es, pues, el producto conmutativo de dos rotaciones de ángulos θ_1 y θ_2 , que actúan sobre planos de \mathbf{R}^4 , ortogonales y suplementarios.

Generadores de $O(n)$ y $SO(n)$

TEOREMA XIII.10.4

|| Sea E un espacio euclídeo de dimensión n . Toda isometría vectorial φ de E es producto de simetrías en número $\leq n$. Si n es **impar**, toda isometría **directa** de E es producto de simetría en número $\leq n - 1$.

Demostración ⁽¹⁾. Tomemos de nuevo las notaciones de XIII.10.2: φ'_k designa la restricción de φ a E_k ($1 \leq k \leq q$) y φ''_l la de φ a F_l ($1 \leq l \leq r$). Además, para todo k ($1 \leq k \leq q$) designamos por E_k^\perp al espacio de dimensión $n - 2$, suma de los subespacios E_m ($m \neq k$) y de los F_l , y E_k^\perp es el ortogonal de E_k . Análogamente, para $1 \leq l \leq r$, el espacio F_l^\perp de dimensión $n - 1$ es la suma de los subespacios F_p ($p \neq l$) y de los E_k , y F_l^\perp es el ortogonal de F_l .

⁽¹⁾ En vista de XIII.10.5 utilizamos aquí un método constructivo. Pero es posible dar una demostración más rápida de este teorema, análoga a la de XIII.11.1.

Designamos por $\overline{\varphi}'_k$ al automorfismo de E cuya restricción a E_k es φ'_k , y cuya restricción a E_k^\perp es la identidad. Finalmente, designamos por $\overline{\varphi}''_l$ al automorfismo de E cuya restricción a F_l es φ''_l , y cuya restricción a F_l^\perp es la identidad. φ es el producto *conmutativo* de las isometrías $\overline{\varphi}'_1, \dots, \overline{\varphi}'_q; \overline{\varphi}''_1, \dots, \overline{\varphi}''_r$.

$\overline{\varphi}''_l$ es la identidad si $\varepsilon_l = 1$, una simetría si $\varepsilon_l = -1$.

Según el corolario de XIII.10.3, φ'_k es el producto de dos simetrías de E_k , o sea

$$(4) \quad \varphi'_k = s_k \circ t_k.$$

Designemos por S_k y T_k a los ejes de las simetrías s_k y t_k , por H_k, L_k a los hiperplanos $S_k + E_k^\perp, T_k + E_k^\perp$ que contienen a E_k^\perp y que pasan respectivamente por S_k y T_k , y por $\overline{s}_k, \overline{t}_k$ las simetrías de hiperplanos H_k, L_k . La relación (4) implica evidentemente

$$(5) \quad \overline{\varphi}'_k = \overline{s}_k \circ \overline{t}_k.$$

De todo ello se deduce el teorema, puesto que φ es el producto de los $\overline{\varphi}'_k$ y de los $\overline{\varphi}''_l$. Si n es cualquiera, el número de las simetrías inducidas es $\leq n$, y si n es impar y si φ es *directa*, una por lo menos de las isometrías $\overline{\varphi}''_l$ es la identidad (cf. nota (1) que sigue a XIII.10.2), por lo tanto, el número de las isometrías introducidas es $\leq n - 1$. c.q.d.

TEOREMA XIII.10.5

|| Sea φ una isometría vectorial **directa** de un espacio euclídeo de dimensión $n \geq 3$; entonces φ es producto de giros en número $\leq n - 1$.

Demostración. Conservemos las notaciones de los teoremas XIII.10.2 y XIII.10.4.

a) Supongamos primeramente que 1 es un valor propio de φ , entonces $\varepsilon_r = 1$. (Esto se realiza ciertamente si n es impar.) Designemos por \overline{u}_r a la simetría respecto del hiperplano F_r^\perp . La relación (5) se puede escribir:

$$\overline{\varphi}'_k = (\overline{s}_k \circ \overline{u}_r) \cdot (\overline{u}_r \circ \overline{t}_k),$$

y $\overline{s}_k \circ \overline{u}_r, \overline{u}_r \circ \overline{t}_k$ son giros.

Por otra parte, puesto que φ es directa, el número de los ε_l ($1 \leq l \leq r - 1$) iguales a -1 , a saber r_1 , es *par*. Cuando $\varepsilon_l = -1$, hacemos $\overline{\varphi}'''_l = \overline{\varphi}''_l \circ \overline{u}_r$; cuando $\varepsilon_l = +1$, hacemos $\overline{\varphi}'''_l = \overline{\varphi}''_l =$ identidad. (Por conmutación de los términos) podemos escribir:

$$\overline{\varphi}'''_1 \circ \overline{\varphi}'''_2 \circ \dots \circ \overline{\varphi}'''_{r-1} = \overline{\varphi}''_1 \circ \overline{\varphi}''_2 \circ \dots \circ \overline{\varphi}''_{r-1} \circ \overline{u}_r^{r_1}.$$

Puesto que r_1 es par, $\overline{u}_r^{r_1}$ es la identidad. Finalmente φ es el producto de los giros $\overline{s}_k \circ \overline{u}_r, \overline{u}_r \circ \overline{t}_k, 1 \leq k \leq q$ y $\overline{\varphi}'''_l$ ($\varepsilon_l = -1$). Su número es $\leq n - 1$. Si n es par, se ve incluso que su número es $\leq n - 2$, puesto que en este caso $r_1 \leq r - 2$.

b) Si 1 no es valor propio de φ , necesariamente n es par. Sea x un elemento $\neq 0$ de E , entonces se tiene $x \neq \varphi(x)$. Dado que $n \geq 3$, existe un hiperplano H que contiene a x y a $\varphi(x)$. Sea Δ el ortogonal en H del vector $x - \varphi(x)$. Δ es un subespacio de dimensión $n - 2$ de E . El giro \bar{u}_Δ , cuyo subespacio de puntos dobles es Δ , cambia x y $\varphi(x)$. Luego, la isometría directa $\theta = \bar{u}_\Delta \circ \varphi$ deja fijo el vector x , y admite 1 como valor propio. En virtud de a), θ es producto de giros en número $\leq n - 2$. Luego $\varphi = \bar{u}_\Delta \circ \theta$ es producto de giros en número $\leq n - 1$.]

§ XIII.11 ISOMETRÍAS AFINES

TEOREMA DE PROLONGACIÓN

(espacio euclídeo E_n)

En el § 9 hemos visto ya que toda isometría afín de E_n está compuesta de una traslación y de una isometría vectorial. Con la ayuda del teorema XIII.9.1, hemos deducido que *toda isometría afín de un espacio euclídeo es una aplicación afín*. (Este resultado lo habríamos podido obtener directamente, como consecuencia de un teorema mucho más general de geometría afín (cf. *Bulletin de l'A.P.M.*, octubre 1970).

El grupo $\mathcal{O}(E_n)$ de las isometrías afines del espacio euclídeo E_n está engendrado por el grupo ortogonal $O(E_n)$ y por el grupo de las traslaciones $\tau(E_n)$.

Elegida una base ortonormal, las *isometrías afines de E_n* son las transformaciones de la forma $\mathcal{Y} = M\mathcal{X} + \mathcal{B}$, en donde $M \in O(n, \mathbf{R})$ (\mathcal{X} e \mathcal{Y} designan las matrices columna de las componentes de x e $y = \varphi(x)$ en (e_i) y \mathcal{B} es una matriz columna dada).

A las transformaciones afines de la forma

$$\mathcal{Y} = M\mathcal{X} + \mathcal{B} \quad (M \in SO(n, \mathbf{R})),$$

se les llama **desplazamientos** o **isometrías directas**, y son las transformaciones afines de E_n cuya parte lineal es una *isometría vectorial directa*. Constituyen un subgrupo de $\mathcal{O}(E_n)$, que designaremos por $\mathcal{D}(E_n)$. Este subgrupo se halla engendrado por $SO(E_n)$ y $\tau(E_n)$, y por lo tanto, es de índice 2 en $\mathcal{O}(E_n)$ (puesto que $SO(E_n)$ es de índice 2 en $O(E_n)$).

Sea φ una isometría afín. El conjunto F de los puntos fijos de φ es evidentemente una variedad afín de E , puesto que las relaciones « $a \in F$, $b \in F$ y c pertenecen a la recta (ab) » implican: $c \in F$. Cuando F es un hiperplano afín, φ es, o la identidad, o bien la simetría respecto a este hiperplano. De hecho, demostraremos la propiedad más general:

TEOREMA XIII.11.1

|| Toda isometría afín φ del espacio euclídeo E_n es producto de simetrías respecto de hiperplanos, en número $\leq n + 1$. La paridad del número de

|| estas simetrías depende únicamente de φ . Con mayor precisión, este número es par o impar según que φ sea o no un desplazamiento.

Demostración. Sea F la variedad afín formada por los puntos fijos de φ , y sea p la dimensión de F (por convenio, si $F = \emptyset$, $p = -1$). Entonces razonaremos por recurrencia descendente sobre p , siendo trivial el resultado para $p = n$.

Hipótesis de recurrencia: si $p < n$ y $\dim(F) \geq p + 1$, φ es producto de simetrías en número $\leq n - p - 1$.

Suponemos entonces que $\dim(F) = p$, y designamos por x a un punto de E_n tal que $x \notin F$. El hiperplano medio H de $[x, \varphi(x)]$ contiene a F , puesto que φ es una isometría. La simetría s_H respecto a H cambia x y $\varphi(x)$. La isometría $\psi = s_H \circ \varphi$ admite como puntos fijos a x y a todos los puntos de F . Por hipótesis de recurrencia, $\psi = s_H \circ \varphi$ es producto de simetrías en número $\leq n - p - 1$. Luego $\varphi = s_H \circ \psi$ es el producto de simetrías en número $\leq n - p$, y ello demuestra la primera parte de nuestro teorema.

Para establecer el final del teorema, es suficiente observar que una simetría respecto a un hiperplano afín está compuesta de una simetría vectorial y de una traslación, por lo tanto, es un elemento de $\mathcal{O}(E_n) \setminus \mathcal{D}(E_n)$. La conclusión proviene del hecho de que $\mathcal{D}(E_n)$ es de índice 2 en $\mathcal{O}(E_n)$. c.q.d.

Teorema de prolongación

Recordemos ante todo algunas propiedades de los sistemas de puntos de un espacio afín.

— A los puntos a_0, a_1, \dots, a_p del espacio afín E_n se les llama *afínmente libres* si, elegido uno de ellos como origen (por ejemplo, a_k), los vectores $\overrightarrow{a_k a_l} = a_l - a_k$ ($k \neq l$) son linealmente independientes (lo cual no depende en absoluto de la elección de a_k). Equivale a decir que la variedad afín engendrada por a_0, a_1, \dots, a_p es de dimensión p .

— Toda parte de un sistema afínmente libre es afínmente libre. Cuando $p = n$, (a_0, a_1, \dots, a_p) es una *base afín* de E_n .

XIII.11.2 Sea (a_0, a_1, \dots, a_n) una base afín del espacio euclídeo E_n . A todo punto x de E_n le hacemos corresponder los $n + 1$ números reales $|a_0 - x|, |a_1 - x|, \dots, |a_n - x|$ (en donde $|a - x|$ designa la distancia euclídea de a a x).
 La aplicación $x \mapsto (|a_0 - x|, \dots, |a_n - x|)$ de E_n en \mathbf{R}^{n+1} es inyectiva. (En otras palabras, un punto de E_n está totalmente determinado si se conocen sus distancias a los puntos a_k .)

Este resultado precisa la noción de «coherencia» de un sistema material, que es la base de la teoría cinemática del sólido.

Demostración. Supongamos que existen x y x' distintos tales que:

$$|a_0 - x| = |a_0 - x'|, \dots, |a_n - x| = |a_n - x'|.$$

El hiperplano medio $[x, x']$ deberá contener entonces a a_0, a_1, \dots, a_n , lo cual es absurdo puesto que su dimensión es $n - 1$. c.q.d.

Es inmediato que una isometría de E_n transforma un sistema afínmente libre en un sistema afínmente libre, puesto que es una aplicación afín biyectiva.

En lo que sigue, vamos a examinar en qué medida una isometría de E_n se puede definir por medio de su restricción a una parte afínmente libre de E_n .

TEOREMA XIII.11.3

|| Sean $A = (a_0, a_1, \dots, a_r)$ una parte afínmente libre del espacio euclídeo E_n , y $\varphi: A \rightarrow E_n$ una aplicación isométrica (e.d. que conserve las distancias). Entonces φ admite una prolongación $\bar{\varphi}$ a E_n , que es una isometría de E_n .

Demostración. Razonaremos por recurrencia sobre p . Para $p = 0$, el teorema significa que el grupo $\mathcal{J}(E_n)$ es transitivo sobre E_n , lo cual es inmediato (por ejemplo, dados dos puntos de E_n , existe una traslación de E que envía el uno al otro).

Supongamos que el teorema es verdadero para el entero p , y sea $A = \{a_0, a_1, \dots, a_{p+1}\}$ una parte afínmente libre de E_n , y $\varphi: A \rightarrow E_n$ una aplicación isométrica. Designemos por ψ a la restricción de φ a $B = (a_0, a_1, \dots, a_p)$. Por hipótesis de recurrencia, ψ se prolonga a una isometría $\bar{\psi}$ de E_n . La aplicación isométrica $\theta = \bar{\psi}^{-1} \circ \varphi$, de A en E_n , deja fijos los elementos a_0, a_1, \dots, a_p . Hagamos $b_{p+1} = \theta(a_{p+1})$. Si $b_{p+1} = a_{p+1}$, θ es la identidad de A , que se prolonga en la identidad de E (pues φ se prolonga a $\bar{\psi}$). Si $b_{p+1} \neq a_{p+1}$, el hiperplano medio H de $[a_{p+1}, b_{p+1}]$ contiene a $\{a_0, a_1, \dots, a_p\}$. Introducimos la simetría s_H respecto a H , y entonces $\bar{\psi} \circ s_H$ es una prolongación de φ . c.q.d.

COROLARIO

|| La imagen de una parte afínmente libre de E_n por una aplicación isométrica es una parte afínmente libre.

Nota. Cuando se dispone de este resultado, se procede más rápidamente a la hora de prolongar la isometría: si x_i ($1 \leq i \leq n$) representan las coordenadas de un punto cualquiera M respecto de una referencia (a_0, a_1, \dots, a_n) de origen a_0 , $\bar{\varphi}(M)$ será el punto con las mismas coordenadas x_i respecto de la referencia

$$(\varphi(a_0), \varphi(a_1), \dots, \varphi(a_n))$$

de origen $\varphi(a_0)$. Pero para saber que $\bar{\varphi}$ así definida es una isometría, es necesario disponer del resultado anterior.

Extensión

TEOREMA XIII.11.4

(Prolongación de las isometrías.)

Sean A una parte **cualquiera** del espacio euclídeo E_n , y $\varphi : A \rightarrow E_n$ una aplicación isométrica. Entonces existe una isometría $\bar{\varphi}$ de E_n que prolonga a φ . Si la variedad afín engendrada por A es E_n , $\bar{\varphi}$ es única. En caso contrario, existe un **desplazamiento** de E_n que prolonga a φ .

Demostración. Sea V la variedad afín engendrada por A , y $p = \dim(V)$. Según el corolario de XIII.11.3 la variedad afín W engendrada por $\varphi(A)$ es de dimensión $\geq p$. Pero como φ es una isometría, es una biyección de A en $\varphi(A)$, por lo que admite una recíproca, que llamaremos ψ , que es una aplicación isométrica de $\varphi(A)$ en A . Aplicando nuevamente el corolario de XIII.11.3 se tiene $\dim(W) \leq p$, luego $\dim(W) = p$.

Sea (a_0, a_1, \dots, a_p) una base afín de V . Según XIII.11.3, la restricción de φ al conjunto $\{a_0, a_1, \dots, a_p\}$ se prolonga a una isometría $\bar{\varphi}$ de E_n . Puesto que $\bar{\varphi}$ es afín, $\bar{\varphi}(V) \subset W$. Sea entonces $x \in A$, se puede escribir:

$$(1) \quad \begin{aligned} |a_0 - x| &= |\varphi(a_0) - \varphi(x)| = |\bar{\varphi}(a_0) - \bar{\varphi}(x)| = |\varphi(a_0) - \bar{\varphi}(x)|, \\ \vdots & \\ |a_p - x| &= |\varphi(a_p) - \varphi(x)| = |\bar{\varphi}(a_p) - \bar{\varphi}(x)| = |\varphi(a_p) - \bar{\varphi}(x)|. \end{aligned}$$

Puesto que $\dim(W) = p$, la familia libre $(\varphi(a_0), \varphi(a_1), \dots, \varphi(a_p))$ es una base afín de W . Además, $\varphi(x) \in W$ y $\bar{\varphi}(x) \in W$. Luego las relaciones (1), teniendo en cuenta XIII.11.2, prueban que $\bar{\varphi}(x) = \varphi(x)$, y esto demuestra la primera parte de nuestro teorema.

Si existen dos isometrías $\bar{\varphi}_1$ y $\bar{\varphi}_2$ que prolonguen a φ , la isometría $\bar{\varphi}_2^{-1} \circ \bar{\varphi}_1$ deja fijo a V . Luego si $p = n$, $\bar{\varphi}_2^{-1} \circ \bar{\varphi}_1$ es la identidad, y $\bar{\varphi}_1 = \bar{\varphi}_2$. Si $p < n$, existen isometrías distintas de la identidad que dejan invariante a V . En particular, existen *simetrías* s que dejan fijo a V . Si $\bar{\varphi}$ no es un desplazamiento, la isometría $s \circ \bar{\varphi}$ es un desplazamiento que, evidentemente, prolonga a φ . c.q.d.

Aplicación a la cinemática del sólido

De XIII.11.4 resulta que la posición de un sólido en E_3 (resp. E_2) está totalmente determinada por la posición de un sistema de puntos del sólido que engendre afínmente E_3 (resp. E_2). Es decir, que contenga por lo menos cuatro puntos no coplanarios (resp. tres puntos no alineados).

Esta cuestión se estudiará más detalladamente en el tomo 3:

Capítulo XIV

Polinomios de varias variables y aplicaciones geométricas

En este capítulo vamos a dar un breve resumen de la difícil teoría de los polinomios con varias variables, fundamentada en la noción de anillo factorial. Este estudio nos permitirá definir de una manera precisa la ecuación de una hipersuperficie algebraica, y, en particular, de una curva algebraica plana. Terminaremos con la noción de curva unicursal, pero no abordaremos el problema de la parametrización de una curva algebraica.

§ XIV.1 ANILLOS FACTORIALES

En lo que sigue, consideraremos únicamente *dominios de integridad*, es decir anillos conmutativos e íntegros. Si A es uno de estos anillos, y si $a \in A$, designaremos por (a) al ideal principal engendrado por a . La relación $(a) = (b)$ es una relación de equivalencia en A^* cuyas clases están formadas por elementos *asociados* dos a dos.

Recordemos que, en un dominio de integridad A , una familia de elementos $(a_i)_{i \in I}$ admite un mcd si (y sólo si) la familia de los ideales (a_i) admite un supremo δ en el conjunto \mathcal{P} de los ideales principales de A , ordenado por inclusión. A todo generador de δ se le llama entonces *un mcd* de los a_i (cf. Cap. III, § 5).

Asimismo, la familia $(a_i)_{i \in I}$ admite un mcm si (y sólo si) la familia de los ideales principales (a_i) admite un ínfimo m en \mathcal{P} . A todo generador de m se le llama entonces un mcm de los a_i .

Si δ existe, todos los mcd de los (a_i) están *asociados*. Asimismo, si m existe, todos los mcm de los (a_i) están asociados.

DEFINICIÓN XIV.1.1

En un dominio de integridad A los elementos de la familia $(a_i)_{i \in I}$ son **primos entre sí** (en conjunto) si su mcd existe y es igual a 1; son **primos** ⁽¹⁾ si el ideal suma $\sum_{i \in I} (a_i)$ es igual a A , es decir, si todo $x \in A$ es una combinación A -lineal de los a_i .

Nota. Si $(a_i)_{i \in I}$ es una familia cualquiera que admite mcd, y si \mathfrak{a} designa el supremo de los ideales (a_i) en \mathcal{P} , se tienen las relaciones de inclusión

$$\sum_{i \in I} (a_i) \subset \mathfrak{a} \subset A.$$

Para que los elementos $(a_i)_{i \in I}$ sean primos entre sí, es suficiente que sean primos. Pero el recíproco es falso, pues la inclusión $\sum_{i \in I} (a_i) \subset \mathfrak{a}$ puede ser estricta, como lo demuestra el siguiente ejemplo:

A es el anillo $K[X, Y]$ (K : cuerpo conmutativo).

El mcd de X e Y es 1, luego X e Y son primos entre sí. Sin embargo, el ideal $(X) + (Y)$ está formado por los polinomios sin término constante, y, por lo tanto, es distinto de A .

DEFINICIÓN XIV.1.2

Un **anillo factorial** A es un dominio de integridad que verifica las siguientes condiciones:

a) Cada elemento no invertible x de A admite una factorización del tipo

$$(1) \quad x = u \prod_{j=1}^n p_j^{\alpha_j},$$

en donde u designa un elemento invertible de A , y p_1, p_2, \dots, p_n elementos irreducibles de A , distintos entre sí, y $\alpha_1, \dots, \alpha_n$ enteros ≥ 1 .

b) Si x admite otra factorización del mismo tipo:

$$x = v \prod_{k=1}^p q_k^{\beta_k},$$

se tiene: $p = n$ y $v = u$; y existe una permutación $\sigma \in \mathfrak{S}_n$ tal que, para $j = 1, 2, \dots, n$, se verifica $\alpha_j = \beta_{\sigma(j)}$, y p_j está asociado a $q_{\sigma(j)}$.

⁽¹⁾ En el texto original se utiliza la palabra «étrangers» para designar tales elementos. (N. del T.).

Ejemplos

- 1) \mathbf{Z} es un anillo factorial; si K es un cuerpo conmutativo, $K[X]$ es un anillo factorial (Cap. IV).
- 2) En general, se puede demostrar que todo anillo principal es factorial.

Divisibilidad en un anillo factorial

En lo que sigue, designaremos por A un dominio de integridad fijo. En $A^* = A \setminus \{0\}$, la relación $(x) = (y)$ ($x \in A^*, y \in A^*$) es una relación de equivalencia, que se traduce por: x e y están asociados. Designemos por \mathcal{C} el conjunto de las clases de equivalencia de los elementos *irreducibles* de A , y por Λ una parte de A que contenga un elemento y sólo uno de cada clase $X \in \mathcal{C}$. Con más precisión, Λ es una parte de A tal que existe una biyección $\varphi : \Lambda \rightarrow \mathcal{C}$ que verifica $x \in \varphi(x)$ para todo $x \in \Lambda$ ⁽¹⁾.

La introducción del conjunto Λ nos evitará el tener que precisar el orden de los factores en los productos de elementos irreducibles.

Según la definición XIV.1.2, decir que A es factorial equivale a la propiedad siguiente:

«Para todo $x \in A^*$, existe un elemento invertible *único* $u(x) \in A$, y una familia *única* de enteros positivos $(\alpha_p(x))_{p \in \Lambda}$, nulos salvo un número finito, tales que

$$(2) \quad x = u(x) \prod_{p \in \Lambda} p^{\alpha_p(x)} . »$$

x es invertible si, y sólo si, $\alpha_p(x) = 0$, para todo $p \in \Lambda$.

XIV.1.1 Sean x e y dos elementos no nulos del anillo factorial A . Con las notaciones anteriores, para que x divida a y , es necesario y suficiente que se verifiquen $\alpha_p(x) \leq \alpha_p(y)$ para todo $p \in \Lambda$.

Demostración. Si $\alpha_p(x) \leq \alpha_p(y)$ para todo $p \in \Lambda$, es claro que, en virtud de (2), x divide a y . Recíprocamente, supongamos que x divide a y . Entonces existe un $z \in A^*$ tal que $y = zx$. De donde, en virtud de (2):

$$y = u(y) \prod_{p \in \Lambda} p^{\alpha_p(y)} = u(x) \cdot u(z) \prod_{p \in \Lambda} p^{\alpha_p(x) + \alpha_p(z)} .$$

⁽¹⁾ En general, la existencia de Λ se desprende del axioma de la elección. Pero, en la práctica, Λ se puede construir. Por ejemplo, si K es un cuerpo conmutativo y si $A = K[X_1, \dots, X_n]$, se tomará como conjunto Λ el conjunto de los polinomios irreducibles tales que el coeficiente de su término de grado máximo, que es el primero en el orden lexicográfico, sea igual a 1.

Dado que la descomposición de y es única, se tiene $u(y) = u(x) u(z)$, y

$$\alpha_p(y) = \alpha_p(x) + \alpha_p(z) \quad \text{para todo } p \in \Lambda, \text{ de donde } \alpha_p(x) \leq \alpha_p(y). \text{ c.q.d.}$$

COROLARIO

En un anillo factorial A , toda familia de elementos $(x_i)_{i \in I}$ de A^ admite un mcd y un mcm.*

Un mcd de los x_i es

$$d = \prod_{p \in \Lambda} p^{\delta_p}, \quad \text{con } \delta_p = \inf_{i \in I} [\alpha_p(x_i)].$$

Un mcm de los x_i es

$$m = \prod_{p \in \Lambda} p^{\mu_i}, \quad \text{con } \mu_i = \sup_{i \in I} [\alpha_p(x_i)],$$

entendiéndose que, si una de las expresiones $\sup_{i \in I} \alpha_p(x_i)$ es igual a $+\infty$, se pone $m = 0$.

Estas fórmulas prueban que las propiedades operatorias del mcd y del mcm dadas en el capítulo IV, § 2, permanecen válidas. Se tiene, por ejemplo,

$$\text{mcd}(\lambda x_i)_{i \in I} = \lambda \text{mcd}(x_i)_{i \in I}, \text{ etc.}$$

TEOREMA XIV.1.2

Si el anillo A es factorial, las relaciones ($a \in A^$, $b \in A^*$, $c \in A^*$, a y b son primos entre sí y a divide a bc) implican (a divide a c). (Teorema de Gauss).*

Demostración. Existe un $x \in A^*$ tal que $bc = ax$. Según (2), se deduce

$$bc = u(b) u(c) \prod_{p \in \Lambda} p^{\alpha_p(b) + \alpha_p(c)} = u(a) u(x) \prod_{p \in \Lambda} p^{\alpha_p(a) + \alpha_p(x)}.$$

La unicidad de la descomposición de bc prueba que $u(b) u(c) = u(a) u(x)$ por un lado, y que

$$(3) \quad \alpha_p(b) + \alpha_p(c) = \alpha_p(a) + \alpha_p(x) \quad \text{para todo } p \in \Lambda.$$

Pero a y b son primos entre sí, lo que significa (corolario anterior), que se cumple $\alpha_p(a) = 0$ o $\alpha_p(b) = 0$ para todo $p \in \Lambda$. En todos los casos, (3) implica pues $\alpha_p(c) \geq \alpha_p(a)$. c.q.d.

Nota. El lector podrá establecer sin ninguna dificultad que, si A es un dominio de integridad en que todo elemento admite una factorización, por lo menos, del tipo (2), y si el teorema de Gauss es válido en A , A es necesariamente factorial; por lo tanto, dicha factorización es única.

TEOREMA XIV.1.3

|| Sean a, b_1, \dots, b_n elementos del anillo factorial A . Si los b_i ($1 \leq i \leq n$) son primos entre sí, dos a dos, y si, para todo i , a es divisible por b_i , entonces a es divisible por el producto $b_1 b_2 \dots b_n$.

Demostración. Por recurrencia sobre n , con la ayuda de XIV.1.1 y XIV.1.2. El resultado es evidente para $n = 1$, y supongamos el teorema verdadero para el orden $n - 1$. Entonces a es divisible por el producto $b_1 b_2 \dots b_{n-1}$, y podemos escribir $a = b_1 \dots b_{n-1} c$, en donde $c \in A$. En virtud del corolario del teorema XIV.1.1, el elemento b_n es primo con el producto $b_1 \dots b_{n-1}$, y divide a a , luego divide a c , según XIV.1.2. Luego a es divisible por $b_1 \dots b_n$, y el teorema es válido también para el orden n . c.q.d.

Ideales primos

DEFINICIÓN XIV.1.3

§ Sea A un anillo conmutativo unífero **cualquiera** El ideal \mathfrak{p} de A es **primo** si el anillo cociente A/\mathfrak{p} es íntegro.

Decir que \mathfrak{p} es un ideal primo significa pues, por una parte, que $\mathfrak{p} \neq A$, y por otra, que las relaciones $x \in A$, $y \in A$, $x \notin \mathfrak{p}$ e $y \notin \mathfrak{p}$, implican $xy \notin \mathfrak{p}$. El ideal $\{0\}$ es primo si, y sólo si, A es íntegro.

Suponemos de nuevo íntegro el anillo A . Si $p \in A^*$ y si el ideal principal (p) es primo, es evidente que p es un elemento irreducible de A . Pero si p es irreducible, el ideal (p) no es necesariamente primo (cf. ejercicios). Este inconveniente desaparece cuando el anillo es factorial:

XIV.1.4 Si el anillo A es factorial, para todo elemento irreducible $p \in A^*$, el || ideal principal (p) es primo.

Demostración. Es una consecuencia inmediata del teorema de Gauss XIV.1.2. ||

§ XIV.2 FACTORIALIDAD DE LOS ANILLOS DE POLINOMIOS

En lo que sigue, A es un anillo factorial fijo. El teorema fundamental es:

TEOREMA XIV.2.1

|| Si A es un anillo factorial, el anillo $A[X]$ es factorial.

El teorema IV.4.4 prueba que este resultado es verdadero si A es un *cuerpo conmutativo*. Su extensión al caso en que A es un *anillo conmutativo* es esencial ya que ello nos permitirá, por paso al anillo $A[X]$, aumentar el número de variables, y establecer la factorialidad de $A[X_1, X_2, \dots, X_n]$. Pero la demostración general de XIV.2.1 exige algunos resultados preliminares que estableceremos primeramente.

● Para todo polinomio $P \in A[X]$, $P \neq 0$, designamos por $\gamma(P)$ al mcd de la familia de los coeficientes no nulos de P . $\gamma(P)$ está definido a menos de un factor invertible de A .

DEFINICIÓN XIV.2.1

§ Un polinomio no nulo $P \in A[X]$ es **primitivo** si $\gamma(P)$ es un elemento invertible de A , lo que equivale a decir que sus coeficientes, en conjunto, son primos entre sí.

Recordemos que los únicos elementos invertibles de $A[X]$ son los elementos invertibles de A (ya que A es íntegro). (Teorema IV.1.5.)

Si P es un polinomio cualquiera $\neq 0$, existe un polinomio P_1 tal que $P = \gamma(P) P_1$. Según la nota que sigue al corolario del teorema XIV.1.1, el polinomio P_1 es primitivo. Recíprocamente, si

$$P = cP_1, \quad c \in A^*,$$

y si P_1 es primitivo, se tiene $c = \gamma(P)$ (salvo para un factor invertible de A).

XIV.2.2 Si P_1 y P_2 son dos polinomios primitivos, $P = P_1 P_2$ es primitivo
|| (Gauss).

Demostración. Vamos a establecer que ningún elemento irreducible p de A puede ser un divisor común a todos los coeficientes de P , lo que demostrará la proposición. Sea $\rho : A \rightarrow A/(p)$ el homomorfismo canónico. ρ se prolonga a un homomorfismo $\bar{\rho} : A[X] \rightarrow A/(p)[X]$ (cf. Cap. IV). Puesto que P_1 es primitivo, $\bar{\rho}(P_1)$ es no nulo. Asimismo, $\bar{\rho}(P_2) \neq 0$. Pero, al ser p irreducible, el anillo $A/(p)$ es íntegro, luego el anillo $A/(p)[X]$ es íntegro (cf. IV.1.3). De todo ello se sigue que $\bar{\rho}(P_1) \bar{\rho}(P_2) = \bar{\rho}(P_1 P_2)$ es no nulo, lo que significa que p no divide a todos los coeficientes de $P_1 P_2$. c.q.d.

XIV.2.3 Para todo par de polinomios no nulos $P, Q \in A[X]$, se tiene (salvo para un factor invertible de A):

$$(1) \quad \gamma(PQ) = \gamma(P) \gamma(Q).$$

Demostración. $P = \gamma(P) P_1$, $Q = \gamma(Q) Q_1$, en donde P_1 y Q_1 son primitivos. De donde: $PQ = \gamma(P)\gamma(Q) P_1 Q_1$, y según XIV.2.2, $P_1 Q_1$ es primitivo. (1) se deduce de lo anterior, si se tienen en cuenta las observaciones que preceden a XIV.2.2.]]

XIV.2.4 Sea K_A el cuerpo de fracciones del anillo factorial A , y sea P un polinomio no constante en $A[X]$. Si existen dos polinomios no constantes $q \in K_A[X]$ y $r \in K_A[X]$ tales que $P = qr$, existen

$$Q \in A[X] \text{ y } R \in A[X], \text{ no constantes,} \\ \text{tales que } P = QR.$$

Demostración. Multiplicamos la relación $P = qr$ por un múltiplo común a los denominadores de todos los coeficientes de q y r , y se deduce en consecuencia la existencia de un $a \in A^*$ y de $S, T \in A[X]$, no constantes, tales que:

$$aP = ST,$$

de donde

$$(2) \quad a\gamma(P) P_1 = \gamma(S) \cdot \gamma(T) S_1 T_1,$$

en donde P_1, S_1 y T_1 son primitivos. Puesto que $S_1 T_1$ es primitivo, se tiene:

$$a\gamma(P) = \gamma(S) \gamma(T),$$

y si se pone $U = a\gamma(P) P_1$, se tiene:

$$\gamma(U) = a\gamma(P) = \gamma(S) \gamma(T),$$

con lo que (2) puede escribirse:

$$aP = a\gamma(P) S_1 T_1, \text{ o sea } P = \gamma(P) S_1 T_1.$$

Por lo tanto, S_1 y T_1 no son constantes, y basta con poner $Q = \gamma(P) S_1$ y $R = T_1$. c.q.d.

Demostración del teorema XIV.2.1

Designemos por K_A el cuerpo de fracciones de A . En virtud de XIV.2.4, un polinomio $p \in A[X]$ es irreducible en $A[X]$ solamente en dos casos: si p es constante en $A[X]$ y se reduce a un elemento irreducible de A , o bien (cuando p es no constante) si p es irreducible en $K_A[X]$ y primitivo en $A[X]$.

Por otra parte, si $q \in K_A[X]$ es un polinomio irreducible, existe un $\lambda \in A^*$ tal que $\lambda q \in A[X]$ y que λq sea primitivo. Entonces se obtiene un conjunto que contiene un elemento y sólo uno de cada clase de elementos irreducibles de $A[X]$ (estas clases son relativas a la relación: x e y están asociados) formando la reunión de los dos conjuntos disjuntos A y M siguientes:

A es un conjunto que contiene un elemento irreducible y sólo uno de cada clase en el anillo A .

M es un conjunto que contiene un polinomio de $A[X]$ irreducible en $K_A[X]$ y primitivo en $A[X]$ y sólo uno de cada clase en $A[X]$. Los $P \in M$ son de grado ≥ 1 .

Vamos a probar que todo polinomio $R \in A[X]$ se puede escribir de manera única en la forma:

$$(3) \quad R = u(R) \left(\prod_{p \in A} p^{\alpha_p(R)} \right) \left(\prod_{P \in M} P^{\beta_P(R)} \right),$$

en donde $u(R)$ es invertible en A , los $\alpha_p(R)$ y los $\beta_P(R)$ son enteros ≥ 0 , todos nulos salvo un número finito.

a) Existencia. Puesto que $K_A[X]$ es factorial (T. IV.4.4), existe un $\zeta \in K_A^*$ y enteros positivos $(\beta_P(R))_{P \in M}$ tales que

$$R = \zeta \prod_{P \in M} P^{\beta_P(R)},$$

siendo todos los $\beta_P(R)$ nulos, salvo un número finito. El polinomio $\prod_{P \in M} P^{\beta_P(R)}$ es primitivo. Razonando como en XIV.2.4, se deduce $\zeta \in A^*$, y por tanto la existencia de un elemento invertible $u(R)$ de A y de una familia $(\alpha_p(R))_{p \in A}$ tales que

$$\zeta = u(R) \prod_{p \in A} p^{\alpha_p(R)}.$$

b) Unicidad. En una relación de la forma (3), el polinomio $S = \prod_{P \in M} P^{\beta_P(R)}$ es primitivo, y el cociente $\frac{R}{S} = u(R) \prod_{p \in A} p^{\alpha_p(R)}$ es un elemento de A . Se tiene, pues, $\gamma(R) = \frac{R}{S}$ salvo para un factor invertible de A , lo que determina $S = \frac{R}{\gamma(R)}$, salvo para un factor invertible de A . Puesto que la factorización de $\frac{R}{\gamma(R)}$ en $K_A[X]$ es única, la familia $(\beta_P(R))$ está determinada de forma única. Luego S está enteramente determinado, así como el elemento $a = \frac{R}{S}$ de A , y se tiene:

$$u(R) \prod_{p \in A} p^{\alpha_p(R)} = a.$$

Puesto que A es factorial, esta descomposición del elemento a es única. c.q.d.

● TEOREMA XIV.2.5

|| Si A es factorial, el anillo $A[X_1, X_2, \dots, X_n]$ es factorial, cualquiera que sea $n \in \mathbf{N}^*$.

Demostración. Por recurrencia sobre n , teniendo en cuenta XIV.2.1, y el isomorfismo canónico entre $A[X_1, \dots, X_n]$ y $A[X_1, \dots, X_{n-1}][X_n]$ (cf. T. IV.7.2).]

Ejemplos:

- 1) Si K es un cuerpo conmutativo, $K[X_1, \dots, X_n]$ es factorial.
- 2) $\mathbf{Z}[X_1, \dots, X_n]$ es factorial.

● TEOREMA XIV.2.6 (fundamental)

|| Sean P, Q_1, \dots, Q_n polinomios con p variables sobre el cuerpo conmutativo K . Si los polinomios Q_i ($1 \leq i \leq n$) son primos entre sí, dos a dos, y si, para todo i , P es divisible por Q_i , entonces P es divisible por el producto $Q_1, Q_2 \dots Q_n$.

Es una consecuencia inmediata de los teoremas XIV.1.3 y XIV.2.5.

Este resultado se ha utilizado en diversas ocasiones, en los ejercicios, a lo largo de esta obra. Citemos, por ejemplo, el cálculo del determinante de Van der Monde y del determinante «circulante» en el § X.3.

Nota. La demostración de XIV.2.1 precisa la forma de los elementos irreducibles de $A[X]$. Más adelante tendremos ocasión de utilizar este resultado.

§ XIV.3 CORRESPONDENCIAS ALGEBRAICAS.
HOMOGRAFÍAS (en característica 0)

En el tomo 3 de esta obra (Geometría) estudiaremos algunas correspondencias algebraicas:

DEFINICIÓN XIV.3.1

Si K designa un cuerpo conmutativo, se llama **correspondencia algebraica** sobre K a toda relación binaria sobre K de la forma

(1)
$$P(x, y) = 0,$$

en donde $P \in K[X, Y]$ es un polinomio no constante tal que $\begin{cases} \text{gr}_X(P) \geq 1 \\ y \\ \text{gr}_Y(P) \geq 1. \end{cases}$

— Cuando K es cualquiera, el grafo de la relación (1) puede ser vacío. Por ejemplo, ello ocurre cuando $K = \mathbf{R}$ y $P = X^2 + Y^2 + 1$. Por este motivo nos limitaremos a los cuerpos algebraicamente cerrados, y muy a menudo a $K = \mathbf{C}$.

Si $x_0 \in K$, el conjunto de los correspondientes a x_0 es el conjunto de los $y \in K$ tales que $P(x_0, y) = 0$, o sea $C(x_0) = \{y \in K \mid P(x_0, y) = 0\}$. El valor x_0 es *regular* si el cardinal de $C(x_0)$ es igual a $q = \text{gr}_Y(P)$, *singular* en caso contrario. Si $\text{card}[C(x_0)] < q$ se dice que x_0 es un valor *singular ordinario*, y si $\text{card}[C(x_0)]$ es infinito, x_0 es un valor *singular degenerado*.

Análogamente se definiría la noción de valor regular o singular de la variable Y .

DEFINICIÓN XIV.3.2

Una **homografía** sobre K es una correspondencia de la forma (1), en la que $\text{gr}_X(P) = \text{gr}_Y(P) = 1$, por lo que se escribirá

$$\alpha xy + \beta x + \gamma y + \delta = 0;$$

Si $\alpha = 0$, se llamará *lineal*.

Propiedades inmediatas

Sea

$$(2) \quad \alpha xy + \beta x + \gamma y + \delta = 0 \quad (\alpha \neq 0)$$

una homografía que no se reduzca a una correspondencia lineal.

Se la puede escribir de la forma: $(\alpha x + \gamma)y + \beta x + \delta = 0$. El valor x_0 es regular si $\alpha x_0 + \gamma \neq 0$; si $\alpha x_0 + \gamma = 0$, x_0 es singular, y únicamente es no degenerado si además se tiene $\beta x_0 + \delta = 0$, por lo tanto, si $\alpha\delta - \beta\gamma = 0$: entonces la homografía se reduce a $(\alpha x + \gamma) \left(y + \frac{\beta}{\alpha} \right) = 0$ y se dice que es *degenerada*. No olvidemos, pues, que la homografía (2) es no degenerada si $\alpha\delta - \beta\gamma \neq 0$, o también si el polinomio $(\alpha X + \gamma)Y + \beta X + \delta$ es *irreducible*, y que entonces no admite valor singular degenerado.

Estos últimos resultados siguen siendo válidos si $\alpha = 0$.

XIV.3.1 Para que la correspondencia algebraica $P(x, y) = 0$ admita a x_0 como valor singular degenerado, es necesario y suficiente que P sea divisible por $X - x_0$.

Demostración. Escribamos

$$P(X, Y) = b_q Y^q + \cdots + b_0, \quad b_i \in K[X], \quad q \geq 1 \text{ y } b_q \neq 0.$$

Para $X = x_0$, el polinomio $P(x_0, Y)$ es nulo para una infinidad de valores de Y , de donde $b_q(x_0) = b_{q-1}(x_0) = \cdots = b_0(x_0) = 0$. Cada b_i es entonces divisible por $X - x_0$. El recíproco es inmediato.]]

DEFINICIÓN XIV.3.3

§ La correspondencia algebraica $P(x, y) = 0$ es **propia** si no admite
 § punto singular degenerado alguno en x o en y .

Podemos ahora demostrar el teorema fundamental que permitió a los «Taurins»⁽¹⁾ de los años treinta hacer razonamientos «sin cálculos»:

TEOREMA XIV.3.2

Sea K un cuerpo algebraicamente cerrado de característica cero, y sea $P(x, y) = 0$ una correspondencia algebraica propia sobre K , que posea las siguientes propiedades:

— existe una parte infinita I de K tal que, para todo $x \in I$, la ecuación $P(x, y) = 0$ posee una raíz y una sola en y ;

— existe una parte infinita J de K tal que, para todo $y \in J$, la ecuación $P(x, y) = 0$ posee una raíz y una sola en x .

Entonces existe un polinomio irreducible $H = \alpha XY + \beta X + \gamma Y + \delta$, de grado 1 en X y en Y , y un entero $h \geq 1$, tales que:

$$P = H^h.$$

Antes de dar la demostración de este teorema, lo expresaremos de una forma más gráfica. Decimos que dos correspondencias algebraicas sobre K son *equivalentes* si tienen el mismo grafo en $K \times K$. El teorema XIV.3.2 implica, en particular, que toda correspondencia algebraica propia sobre K , que posea las propiedades indicadas, sea *equivalente a una homografía no degenerada*.

Demostración del teorema XIV.3.2

a) Supongamos primeramente que el polinomio $P(X, Y)$ es irreducible. Escribamos

$$P = a_p X^p + \cdots + a_0, \quad (a_k \in K[Y], \quad p \geq 1, \quad a_p \neq 0).$$

⁽¹⁾ Candidato a la Ecole polytechnique. (N. del T.).

Vamos a demostrar que $p = 1$. Si tuviéramos que $p \geq 2$, el polinomio P'_x tendría grado $p - 1 \geq 1$ en X (ya que K es de característica 0). Sea $\Delta(Y)$ el discriminante de P , considerado como polinomio en X , con coeficientes en $K[Y]$. $\Delta(Y)$ que es la resultante de P y P'_x (cf. Cap. VI) es un polinomio en Y con coeficientes en K .

Si tuviéramos $\Delta(Y) = 0$, existirían polinomios U y $V \in K[Y][X]$ tales que

$$(3) \quad UP + VP'_x = 0, \quad \text{gr}_X(U) < \text{gr}_X(P'_x) \quad \text{y} \quad \text{gr}_X(V) < \text{gr}_X(P) = p \quad U \neq 0$$

y $V \neq 0$ (cf. T. VI.1.1).

El examen de los grados que intervienen en (3) muestra que P y P'_x no serían primos entre sí en $L[X]$, en donde L designa el cuerpo $K(Y)$. Puesto que P'_x no es constante en $L[X]$, tendríamos que P no sería irreducible en $L[X]$. Aplicando XIV.2.4, con $A = K[Y]$ y $K_A = K(Y)$, vemos que P no sería irreducible en $K[Y][X] = K[X, Y]$, en contra de la hipótesis.

De todo ello se deduce que $\Delta(Y) \neq 0$. Existe, pues, un $y_0 \in J$ tal que $\Delta(y_0) \neq 0$, y la ecuación $P(x, y_0) = 0$ tiene sus p raíces en x distintas, y puesto que $p \geq 2$, estaríamos en contradicción con las hipótesis de XIV.3.2. Se deduce, pues, que $p = 1$, y análogamente se demuestra que el grado de P en Y es 1 ⁽¹⁾.

b) Si P es cualquiera, lo descomponemos en factores irreducibles en $K[X, Y]$, con lo cual $P = P_1 P_2 \dots P_m$, y dado que la correspondencia es propia, el grado de cada P_i en X y en Y es ≥ 1 .

Pongamos

$$P_i = a_{i,p_i} X^{p_i} + \dots + a_{i,0} = b_{i,q_i} Y^{q_i} + \dots + b_{i,0},$$

en donde

$$p_i \geq 1, \quad q_i \geq 1, \quad a_{i,k} \in K[Y], \quad b_{i,k} \in K[X], \quad a_{i,p_i} \neq 0, \quad b_{i,q_i} \neq 0 \quad (1 \leq i \leq m).$$

Designemos por E y por F a los conjuntos de las raíces de todos los b_{i,q_i} y todos los a_{i,p_i} ($1 \leq i \leq m$). Pongamos

$$I_1 = I \setminus E, \quad J_1 = J \setminus F.$$

Para cada $x_0 \in I_1$, $P_i(x_0, y) = 0$ admite, por lo menos, una raíz en y (puesto que $P_i(x_0, y)$ no es constante); y esta raíz es única (ya que es raíz de $P(x_0, y) = 0$). Asimismo, para cada $y_0 \in J_1$, $P_i(x, y_0)$ admite una sola raíz en x . Por lo tanto, cada uno de los P_i verifica las hipótesis de XIV.3.2, en donde se ha reemplazado I por I_1 y J por J_1 (I_1 y J_1 son infinitos, ya que I y J son infinitos y E y F son finitos). En virtud de a), existen entonces

$$\alpha_i, \beta_i, \gamma_i, \delta_i \in K \quad (1 \leq i \leq m)$$

(1) Más adelante daremos una interpretación geométrica de esta demostración.

tales que

$$P_i = \alpha_i XY + \beta_i X + \gamma_i Y + \delta_i,$$

siendo el grado de P_i en X y en Y igual a 1. De donde:

$$P = \prod_{i=1}^m (\alpha_i XY + \beta_i X + \gamma_i Y + \delta_i).$$

En esta situación es fácil ver que todos los P_i son proporcionales (es decir, asociados). En efecto, las hipótesis hechas sobre P implican que, para $i \neq j$, existe una infinidad de valores x_0 tales que las ecuaciones $P_i(x_0, y) = 0$ y $P_j(x_0, y) = 0$ tienen la misma raíz en y (de lo contrario, dejando de lado el caso en que x_0 es un valor singular de P_i o de P_j , la ecuación $P(x_0, y) = 0$ tendría dos raíces en y). Dicho de otra manera: existiría una infinidad de valores de x tales que

$$\frac{\alpha_i x + \gamma_i}{\beta_i x + \delta_i} = \frac{\alpha_j x + \gamma_j}{\beta_j x + \delta_j},$$

lo que exigiría $\alpha_i \beta_j - \beta_i \alpha_j = 0$, $\gamma_i \delta_j - \delta_i \gamma_j = 0$. Intercambiando X e Y , tendríamos asimismo:

$$\alpha_i \gamma_j - \alpha_j \gamma_i = \beta_i \delta_j - \beta_j \delta_i = 0,$$

de donde resulta nuestra afirmación. A fin de cuentas, existe un $\lambda \in K^*$ y $\alpha, \beta, \gamma, \delta \in K$, tales que $\alpha XY + \beta X + \gamma Y + \delta$ es irreducible y de grado 1 en X e Y , y tal que

$$P = \lambda(\alpha XY + \beta X + \gamma Y + \delta)^m. \text{ c.q.d.}$$

Nota. El razonamiento de a) demuestra, de hecho, la siguiente proposición: si $P(x, y) = 0$ es una correspondencia algebraica propia y si P es irreducible, el grado de P en X es exactamente igual a $p(p \geq 1)$, y su grado respecto de Y es exactamente igual a $q(q \geq 1)$ cuando se presentan las condiciones siguientes:

- existe una parte infinita I de K tal que todo $x \in I$ tiene exactamente q correspondientes;
- existe una parte infinita J de K tal que todo $y \in J$ tiene exactamente p correspondientes.

§ XIV.4 HIPERSUPERFICIES ALGEBRAICAS EN \mathbf{C}^n ($n \geq 2$)

DEFINICIÓN XIV.4.1

$\left\{ \begin{array}{l} \text{Sea } P \text{ un polinomio no constante de } \mathbf{C}[X_1, \dots, X_n]. \text{ Se llama } \mathbf{hiper-} \\ \text{superficie definida por } P \text{ al conjunto de los puntos } x = (x_1, \dots, x_n) \end{array} \right\}$

$\{ \in \mathbf{C}^n \text{ tales que } P(x) = 0. \text{ Si } \mathcal{H} \text{ es una hipersuperficie, se llama ciclo ligado a } \mathcal{H} \text{ a todo polinomio } P \text{ tal que la hipersuperficie definida por } P \text{ sea } \mathcal{H}. \}$

Si $P \in \mathbf{C}[X_1, \dots, X_n]$ es no constante, la hipersuperficie \mathcal{H}_P definida por P es no vacía. En efecto, existe un $i \in \mathbf{N}_n^*$ tal que P es de grado ≥ 1 en X_i . Por lo tanto, se tiene:

$$P = a_p X_i^p + \dots + a_0,$$

con $p \geq 1$, $a_k \in \mathbf{C}[X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n]$ y $a_p \neq 0$.

Existen, pues, $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ tales que

$$a_p(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \neq 0.$$

El polinomio $F(X_i) = P(x_1, \dots, x_{i-1}, X_i, x_{i+1}, \dots, x_n)$ es de grado p , y puesto que \mathbf{C} es algebraicamente cerrado, existe un $x_i \in \mathbf{C}$ tal que $F(x_i) = 0$, de donde $(x_1, \dots, x_n) \in \mathcal{H}_P$. Finalmente, existe un único ciclo ligado a $\mathcal{H} = \mathbf{C}^n$, y el polinomio nulo (cf. T. IV.7.5).

Polinomios regulares respecto a una variable

Consideremos una matriz invertible $A = [a_{ij}]_{1 \leq i \leq n, 1 \leq j \leq n}$ con elementos complejos, y hagamos $X_i = \sum_j a_{ij} Y_j$, en donde los Y_j son nuevas variables. Si $P \in \mathbf{C}[X_1, \dots, X_n]$, designemos por $\varphi_A(P)$ al polinomio

$$P\left(\sum_j a_{1j} Y_j, \sum_j a_{2j} Y_j, \dots, \sum_j a_{nj} Y_j\right).$$

Si (e_1, \dots, e_n) es la base canónica de \mathbf{C}^n , y (f_1, \dots, f_n) es la referencia de \mathbf{C}^n cuya matriz respecto de (e_i) es A , los puntos de la hipersuperficie definida por P son los puntos cuyas coordenadas (y_1, \dots, y_n) en (f_i) verifican

$$\varphi_A(P)(y_1, \dots, y_n) = 0.$$

Luego la hipersuperficie definida por P en (e_i) es también la definida por $\varphi_A(P)$ en (f_j) . Además, evidentemente P es irreducible si, y sólo si, $\varphi_A(P)$ es irreducible, y P y $\varphi_A(P)$ tienen el mismo grado total. Finalmente, un polinomio P divide a un polinomio Q si, y sólo si, $\varphi_A(P)$ divide a $\varphi_A(Q)$ ⁽¹⁾.

⁽¹⁾ Con precisión, la aplicación $P \mapsto \varphi_A(P)$ es un automorfismo de la \mathbf{C} -álgebra $\mathbf{C}[X_1, \dots, X_n]$.

DEFINICIÓN XIV.4.2

§ A un polinomio $P \in \mathbf{C}[X_1, \dots, X_n]$ de grado total $p \geq 1$, se le llama **regular en X_i** si se puede escribir:

$$P = a_p X_i^p + a_{p-1} X_i^{p-1} + \dots + a_0, \quad \text{con } a_p \in \mathbf{C}^*,$$

§ y $a_k \in \mathbf{C}[X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n]$ para $0 \leq k \leq p-1$.

XIV.4.1 Sean P_1, \dots, P_m m polinomios con n variables, de grados totales p_1, \dots, p_m ($p_i \geq 1$). Existe una matriz invertible $A \in M_n(\mathbf{C})$ tal que todos los polinomios $\varphi_A(P_i)$ son regulares en X_1 .

Demostración. Descompongamos cada P_i en suma de polinomios homogéneos de grados $p_i, p_i - 1, \dots, 0$:

$$P_i = B_{i,p_i} + B_{i,p_i-1} + \dots + B_{i,0} \quad (B_{i,p_i} \neq 0 \text{ y homogéneo de grado } p_i).$$

Existe un $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbf{C}^n$ tal que $B(\alpha_1, \alpha_2, \dots, \alpha_n) \neq 0$, en donde

$$B = \prod_{i=1}^m B_{i,p_i},$$

y existe una matriz invertible $A = [a_{ij}] \in M_n(\mathbf{C})$ tal que

$$A \times \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Para todo i , se tendrá $\varphi_A(P_i) = \varphi_A(B_{i,p_i}) + \dots + \varphi_A(B_{i,0})$, y $\varphi_A(B_{i,p_i})$ es la parte homogénea de grado p_i de $\varphi_A(P_i)$. Por construcción, se tiene:

$$\varphi_A(B_{i,p_i})(1, 0, \dots, 0) = B_{i,p_i}(\alpha_1, \dots, \alpha_n) \neq 0,$$

lo que significa que $\varphi_A(P_i)$ es regular en X_1 . c.q.d.

XIV.4.2 Sean P y $Q \in \mathbf{C}[X_1, \dots, X_n]$ dos polinomios no constantes, tales que el ideal $(P) + (Q)$ engendrado por P y Q sea distinto de $\mathbf{C}[X_1, \dots, X_n]$. Las hipersuperficies ligadas a P y Q , a saber \mathcal{H}_P y \mathcal{H}_Q , tienen, por lo menos, un punto común.

Demostración. En virtud de XIV.4.1 y según la observación que precede a la definición XIV.4.2, podemos suponer P y Q regulares en X_1 , de modo que

$$P = \alpha_p X_1^p + A_{p-1} X_1^{p-1} + \cdots + A_0$$

$$Q = \beta_q X_1^q + B_{q-1} X_1^{q-1} + \cdots + B_0, \quad p \text{ y } q \geq 1, \alpha_p, \beta_q \in \mathbf{C}^*,$$

$$A_i, B_j \in \mathbf{C}[X_2, \dots, X_n].$$

Sea $R[X_2, \dots, X_n]$ la *resultante* de P y Q considerados como polinomios en X_1 . Existen $U, V \in \mathbf{C}[X_2, \dots, X_n][X_1]$ tales que

$$(1) \quad UP + VQ = R[X_2, \dots, X_n].$$

Puesto que el ideal $(P) + (Q)$ es propio, de (1) se sigue que si R es constante, es R nulo. En todos los casos, existen pues

$$x_2, \dots, x_n \in \mathbf{C} \text{ tales que } R(x_2, \dots, x_n) = 0.$$

Pongamos $F(X_1) = P(X_1, x_2, \dots, x_n)$ y $G(X_1) = Q(X_1, x_2, \dots, x_n)$. Puesto que

$$R(x_2, \dots, x_n) = 0,$$

existe x_1 tal que $F(x_1) = G(x_1) = 0$, y el punto $x = (x_1, \dots, x_n)$ pertenece a $\mathcal{H}_P \cap \mathcal{H}_Q$. c.q.d.

TEOREMA XIV.4.3

|| Sea $P \in \mathbf{C}[X_1, \dots, X_n]$ un polinomio no constante, que defina la hipersuperficie \mathcal{H}_P . Sea H un polinomio no nulo tal que la relación $x \in \mathcal{H}_P$ implique $H(x) = 0$. Entonces existe un $\rho \in \mathbf{N}^*$ tal que P divide a H^ρ .

Demostración. En \mathbf{C}^{n+1} , designamos por \mathcal{L}_P y por \mathcal{L}_H las hipersuperficies definidas por los polinomios $P(X_1, \dots, X_n)$ y $1 - T.H(X_1, \dots, X_n)$ de $\mathbf{C}[X_1, \dots, X_n; T]$ (en donde T designa una nueva variable). Según las hipótesis, \mathcal{L}_P y \mathcal{L}_H no tienen ningún punto en común. Entonces XIV.4.2 implica la existencia de U y $V \in \mathbf{C}[X_1, \dots, X_n; T]$ tales que

$$(2) \quad UP + V(1 - T.H) = 1.$$

Reemplacemos T por el elemento $\frac{1}{H}$ del cuerpo $\mathbf{C}(X_1, \dots, X_n)$. Si ρ designa el grado de U en T , (2) implica

$$H^\rho U\left(X_1, \dots, X_n; \frac{1}{H}\right) \cdot P(X_1, \dots, X_n) = H^\rho;$$

y puesto que $H^p U\left(X_1, \dots, X_n; \frac{1}{H}\right)$ es un polinomio en X_1, \dots, X_n , el teorema queda demostrado.]]

● COROLARIO

|| Sea $P \in \mathbf{C}[X_1, \dots, X_n]$ un polinomio **irreducible**. Si H es un polinomio $\neq 0$ tal que la relación $x \in \mathcal{H}_P$ implica $H(x) = 0$, H es un múltiplo de P .

Demostración. Existe un $\rho \in \mathbf{N}^*$ tal que P divide a H^ρ . En virtud de la factorialidad de $\mathbf{C}[X_1, \dots, X_n]$ (cf. § 1), P divide a H . c.q.d.

Este resultado ha sido utilizado ya en algunos razonamientos realizados en el capítulo VI. El teorema XIV.4.3 y su corolario son casos particulares del *teorema de los ceros de Hilbert*.

Vamos a precisar estos resultados.

XIV.4.4 Sean $P, F \in \mathbf{C}[X_1, \dots, X_n]$, $P \neq 0$. Si las relaciones

|| $(a_1, \dots, a_n) \in \mathbf{C}^n$ y $P(a_1, \dots, a_n) \neq 0$
|| implican $F(a_1, \dots, a_n) = 0$, el polinomio F es nulo.

Demostración. Según IV.5.2, es suficiente demostrar que la función polinomio asociada a F sobre \mathbf{C}^n es nula. Sea (b_1, \dots, b_n) un punto que no pertenezca a la hipersuperficie \mathcal{H}_P definida por P . (Tal punto existe puesto que $P \neq 0$). Vamos a demostrar que, para todo punto $(a_1, \dots, a_n) \in \mathcal{H}_P$, el polinomio

$$\varphi(t) = F(a_1 + t(b_1 - a_1), \dots, a_n + t(b_n - a_n))$$

admite una infinidad de raíces. A este fin, designamos por $M(t)$ ($t \in \mathbf{C}$) al punto $(a_i + t(b_i - a_i))_{1 \leq i \leq n}$. El polinomio

$$\psi(t) = P(a_1 + t(b_1 - a_1), \dots, a_n + t(b_n - a_n))$$

no es nulo, pues $\psi(1) = P(b_1, \dots, b_n) \neq 0$. Por lo tanto, admite un número finito de raíces, que designaremos por t_1, \dots, t_p . Para

$$t \neq t_1, t_2, \dots, t_p,$$

el punto $M(t)$ no pertenece, pues, a \mathcal{H}_P , de donde $\varphi(t) = F(M(t)) = 0$ en virtud de las hipótesis, lo que prueba nuestra afirmación. Puesto que posee una infinidad de raíces, el polinomio φ es nulo. Luego, en particular, $\varphi(0) = F(a_1, \dots, a_n) = 0$.]]

Nota

Es claro que XIV.4.4 permanece válido (pues su demostración se extiende sin modificación) cuando se substituye \mathbf{C} por un anillo K unífero, íntegro e infinito.

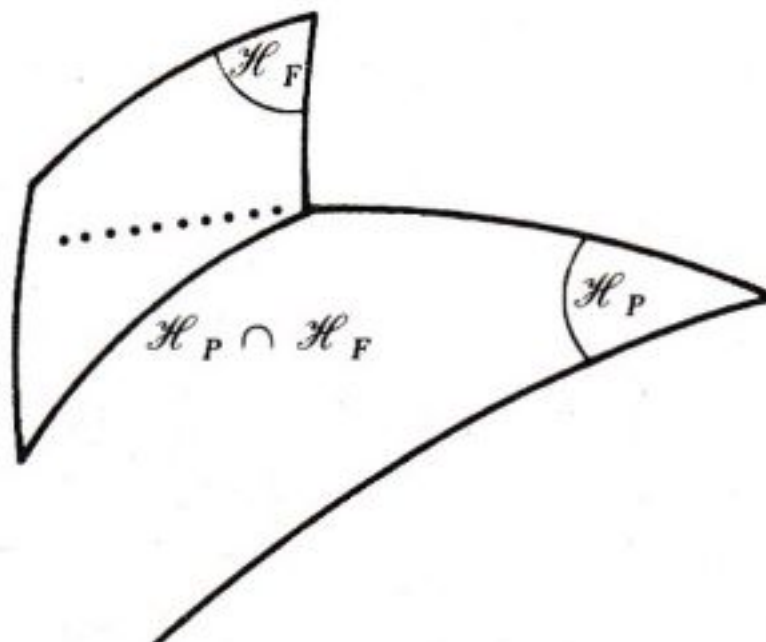
TEOREMA XIV.4.5

Sean $P \in \mathbf{C}[X_1, \dots, X_n]$ un polinomio **irreducible** y $F \in \mathbf{C}[X_1, \dots, X_n]$ un polinomio no divisible por P . Designemos por \mathcal{H}_P y \mathcal{H}_F a las hipersuperficies asociadas, respectivamente, a P y a F . Finalmente, sea G un polinomio cualquiera. Si la relación $(a_1, \dots, a_n) \in \mathcal{H}_P \setminus \mathcal{H}_F$ implica $G(a_1, \dots, a_n) = 0$, entonces la relación

$$(a_1, \dots, a_n) \in \mathcal{H}_P \text{ implica } G(a_1, \dots, a_n) = 0.$$

En consecuencia, G es un múltiplo de P .

Comentario. Este teorema significa que si F no es divisible por P , $\mathcal{H}_P \cap \mathcal{H}_F$ es una parte «despreciable» de \mathcal{H}_P , en el sentido de que un polinomio G que se anule sobre $\mathcal{H}_P \setminus (\mathcal{H}_P \cap \mathcal{H}_F)$, se anula necesariamente sobre toda la hipersuperficie \mathcal{H}_P .



Demostración de XIV.4.5. El polinomio $H = GF$ es tal que la relación

$$(a_1, \dots, a_n) \in \mathcal{H}_P \text{ implica } H(a_1, \dots, a_n) = 0.$$

Según el corolario de XIV.4.3, H es múltiplo de P . Puesto que, por hipótesis, F y P son primos entre sí, resulta del teorema de Gauss que G es múltiplo de P , lo que demuestra las afirmaciones de XIV.4.5. c.q.d.

Ciclos ligados a una hipersuperficie

Si la hipersuperficie \mathcal{H}_F está definida por el polinomio F , descomponemos F en factores irreducibles distintos:

$$(3) \quad F = P_1^{\alpha_1} P_2^{\alpha_2} \dots P_m^{\alpha_m}, \quad \alpha_i \geq 1 \text{ para } 1 \leq i \leq m.$$

Sea \mathcal{H}_p la hipersuperficie definida por P_p , y es evidente que \mathcal{H}_F es la reunión de las \mathcal{H}_p ($1 \leq p \leq m$). A las \mathcal{H}_p se les llama *componentes irreducibles* de \mathcal{H}_F . Los resultados de este § nos permiten establecer el teorema siguiente:

TEOREMA XIV.4.6

Si \mathcal{H}_F es la hipersuperficie definida por el polinomio F de la relación (3), todo ciclo ligado a \mathcal{H}_F es de la forma

$$P_1^{\beta_1} P_2^{\beta_2} \dots P_m^{\beta_m}, \quad \text{con } \beta_i \geq 1 \text{ para } 1 \leq i \leq m \quad (1).$$

Demostración. Sea G un ciclo ligado a \mathcal{H}_F . Según XIV.4.3, existe $\rho \geq 1$ entero, tal que G^ρ es un múltiplo de F . Escribamos

$$(4) \quad G^\rho = Q P_1^{\alpha_1} P_2^{\alpha_2} \dots P_m^{\alpha_m}, \quad Q \in \mathbf{C}[X_1, \dots, X_n].$$

Por hipótesis, la hipersuperficie definida por G es \mathcal{H}_F . Según XIV.4.3, existe también $\sigma \geq 1$ entero, tal que F^σ es múltiplo de G , o sea:

$$(5) \quad F^\sigma = MG, \quad (M \in \mathbf{C}[X_1, \dots, X_n]).$$

De (5) se deduce que los únicos factores irreducibles de G posibles son los P_i ; y (4) muestra entonces que cada uno de ellos figura en la descomposición de G con una potencia ≥ 1 . c.q.d.

El teorema XIV.4.6 precisa lo que debemos entender por «ecuación» de una hipersuperficie. Se observa que no es posible establecer un ligamen entre esta noción (correcta) de ecuación, que acabamos de dar, y la noción ingenua de «ecuación de una hipersuperficie», salvo en casos muy particulares. Se tiene, por ejemplo, el

COROLARIO

Sea \mathcal{H}_F la hipersuperficie definida por el polinomio F . Si se pone $d = \text{gr}(F)$, y si F se descompone de la forma siguiente:

$$F = P_1 P_2 \dots P_m,$$

en factores irreducibles distintos, todo ciclo de grado d ligado a \mathcal{H}_F es de la forma λF ($\lambda \in \mathbf{C}^*$).

(1) Recíprocamente, es evidente que todo polinomio de esta forma es un ciclo ligado a \mathcal{H}_F .

§ XIV.5 CURVAS ALGEBRAICAS Y CURVAS ALGEBRAICAS UNICURSALES EN \mathbf{C}^2

Por definición, una *curva algebraica* en \mathbf{C}^2 es una hipersuperficie algebraica de \mathbf{C}^2 . Empezaremos con el teorema siguiente, que constituye una introducción al teorema general de Bezout.

XIV.5.1 Sean P y $Q \in \mathbf{C}[X, Y]$ dos polinomios **irreducibles** y no asociados.
 || El conjunto de los puntos comunes a las curvas algebraicas Γ_P y Γ_Q , definidas por P y Q , es **finito**.

Demostración. Según XIV.4.1, podemos suponer que P y Q son regulares en X . De donde, poniendo $m = \text{gr}(P)$, $n = \text{gr}(Q)$:

$$\begin{aligned} P &= X^m + A_{m-1} X^{m-1} + \cdots + A_0 & A_k &\in \mathbf{C}[Y], \\ Q &= X^n + B_{n-1} X^{n-1} + \cdots + B_0 & B_k &\in \mathbf{C}[Y]. \end{aligned}$$

Sea $R(Y)$ el polinomio en Y resultante de P y Q , considerados como polinomios en X . Si $R = 0$, deduciríamos la existencia de $U \neq 0$ y $V \neq 0$ tales que $UP = -VQ$ (cf. T. VI.1.1), con $U, V \in \mathbf{C}(Y)[X]$. Por consideraciones sobre el grado, P y Q no serían primos entre sí en $L[X]$, en donde L designa el cuerpo $\mathbf{C}(Y)$.

Entonces serían posibles dos casos:

Primer caso: P y Q no están asociados en $L[X]$. De ello se deduciría que uno de ellos, por lo menos, por ejemplo P , no sería irreducible en $L[X]$. Pero entonces, según XIV.2.4, P no sería irreducible en $\mathbf{C}[X, Y]$ identificado con $\mathbf{C}[Y][X]$, contrariamente a la hipótesis.

Segundo caso: P y Q están asociados en $L[X]$. Existiría entonces una fracción racional no nula $S \in \mathbf{C}(Y)$, por ejemplo, $S = \frac{A}{B}$, en donde $A, B \in \mathbf{C}[Y]$, tal que

$$P = SQ, \text{ de donde } BP = AQ.$$

El teorema de Gauss implicaría entonces que $P \mid A$ y $Q \mid B$ en $\mathbf{C}[X, Y]$, lo cual es absurdo (puesto que A y B no dependen de X).

De todo ello se concluye que $R(Y) \neq 0$, y el número de raíces de R es, pues, **finito**; pero las raíces de R son las ordenadas de los puntos de $\Gamma_P \cap \Gamma_Q$. Se demuestra análogamente que el número de abscisas de estos puntos es **finito**. ||

COROLARIO

|| Si F y $G \in [X, Y]$ son polinomios **primos entre sí**, y si \mathcal{L}_F y \mathcal{L}_G son las curvas algebraicas definidas por F y G , el conjunto $\mathcal{L}_F \cap \mathcal{L}_G$ es **finito**.

Demostración. Sean $\Gamma_1, \dots, \Gamma_r$ las componentes irreducibles de \mathcal{L}_F , $\Delta_1, \dots, \Delta_s$ las de \mathcal{L}_G . Según XIV.5.1, para todo i y todo j , $\Gamma_i \cap \Delta_j$, es finito. Luego $\mathcal{L}_F \cap \mathcal{L}_G = \bigcup_{i,j} (\Gamma_i \cap \Delta_j)$ es finito. c.q.d.

Aplicaciones racionales y curvas unicursales

Después de haber definido las curvas algebraicas planas por una ecuación (curvas definidas implícitamente), vamos a estudiar otro modo de definición, que se inspira en la noción de trayectoria (curvas paramétricas).

DEFINICIÓN XIV.5.1

Se llama **aplicación racional** de una variable (o parametrización racional en una variable), con valores en \mathbf{C}^2 , a un par de fracciones racionales en una variable $F, G \in \mathbf{C}(T)$. Si Ω es la reunión de los conjuntos de polos de F y G , $\mathbf{C} \setminus \Omega$ es el **dominio de definición** de la aplicación. Para todo $t \in \mathbf{C} \setminus \Omega$, el punto de coordenadas $(F(t), G(t))$ en \mathbf{C}^2 es el **punto correspondiente al valor t del parámetro**.

Sea $M(t)$ el punto $(F(t), G(t))$. Si $(F_1(t), G_1(t))$ son las coordenadas de $M(t)$ en otra base de \mathbf{C}^2 , $F_1(t)$ y $G_1(t)$ son fracciones racionales, y el dominio de definición común a F_1 y G_1 es el mismo que el común a F y G .

Un punto $M_0 \in \mathbf{C}^2$ es **múltiplo** si existen t_1 y $t_0 \in \mathbf{C} \setminus \Omega$, distintos, tales que $M(t_0) = M(t_1) = M_0$.

DEFINICIÓN XIV.5.2

Se llama aplicación (o parametrización) racional **propia de \mathbf{C}^2** a toda aplicación racional con valores en \mathbf{C}^2 , que sólo posea un número finito de puntos múltiples.

DEFINICIÓN XIV.5.3

Dos aplicaciones racionales propias $(t, M(t))$, $(u, N(u))$ son **equivalentes** si existen dos partes finitas I y J de \mathbf{C} y una parte Γ de \mathbf{C}^2 tales que las dos aplicaciones

$$\left\{ \begin{array}{l} t \mapsto M(t) \\ \mathbf{C} \setminus I \rightarrow \Gamma \end{array} \right. \quad y \quad \left\{ \begin{array}{l} u \mapsto N(u) \\ \mathbf{C} \setminus J \rightarrow \Gamma \end{array} \right.$$

sean biyectivas

La equivalencia así definida es una relación de equivalencia en el conjunto de las aplicaciones racionales propias.

TEOREMA XIV.5.2

Si $M(t)$ y $N(u)$ son dos aplicaciones racionales propias equivalentes, existe una homografía no degenerada

$$H(t, u) = \alpha tu + \beta t + \gamma u + \delta,$$

tal que las relaciones $(M(t)$ y $N(u)$ están definidas y $M(t) = N(u))$ implican $H(t, u) = 0$.

Este resultado es el que se expresa, de forma intuitiva, cuando se dice que dos parámetros distintos de una misma curva unicursal se deducen uno de otro por un cambio homográfico de coordenadas.

Demostración. Sean $F(t) = \frac{A(t)}{B(t)}$, $G(t) = \frac{C(t)}{D(t)}$ las formas irreducibles de las componentes de $M(t)$, y sean asimismo $U(u) = \frac{P(u)}{Q(u)}$, $V(u) = \frac{R(u)}{S(u)}$ las de las componentes de $N(u)$ (A, B, C, D, P, Q, R, S pertenecen a $\mathbf{C}[T]$).

Designemos por I y J partes finitas de \mathbf{C} y por Γ una parte de \mathbf{C}^2 tales que las aplicaciones

$$\left\{ \begin{array}{l} t \mapsto M(t) \\ \mathbf{C} \setminus I \rightarrow \Gamma \end{array} \right. \quad \text{y} \quad \left\{ \begin{array}{l} u \mapsto N(u) \\ \mathbf{C} \setminus J \rightarrow \Gamma \end{array} \right.$$

sean biyecciones.

Para $t \in \mathbf{C} \setminus I$ y $u \in \mathbf{C} \setminus J$, la relación $M(t) = N(u)$ equivale a:

$$\text{y} \quad \left\{ \begin{array}{l} A(t)Q(u) - B(t)P(u) = 0 \\ C(t)S(u) - D(t)R(u) = 0 \end{array} \right\},$$

o sea

$$(1) \quad \left\{ \begin{array}{l} \varphi(t, u) = 0 \\ \psi(t, u) = 0 \end{array} \right\},$$

en donde φ y ψ son polinomios. Si uno de estos polinomios, por ejemplo φ , fuera nulo; significaría que $F(t)$ y $U(u)$ se reducen a constantes. Según las hipótesis, esto no puede ocurrir para φ y ψ simultáneamente.

Designemos por Δ el mcd de φ y ψ en $\mathbf{C}[X, Y]$. Entonces se tiene:

$$\varphi = \Delta\varphi_1, \quad \psi = \Delta\psi_1,$$

en donde φ_1 y ψ_1 son primos entre sí ⁽¹⁾. Según el corolario XIV.5.1, el conjunto de los pares (t, u) tales que $\varphi_1(t, u) = \psi_1(t, u) = 0$ es finito. Podemos, pues, hallar una parte finita K de \mathbf{C} tal que las relaciones

$$t \in \mathbf{C} \setminus K, \quad u \in \mathbf{C} \setminus K \quad \text{y} \quad M(t) = N(u)$$

equivalgan a:

$$t \in \mathbf{C} \setminus K, \quad u \in \mathbf{C} \setminus K \quad \text{y} \quad \Delta(t, u) = 0,$$

y que las aplicaciones $t \mapsto M(t)$ y $u \mapsto N(u)$ sean biyecciones de $\mathbf{C} \setminus K$ en Γ . Es inmediato que la correspondencia $\Delta(t, u) = 0$ verifica las hipótesis de XIV.3.2, y por tanto, en virtud del teorema XIV.3.2, se deduce el resultado. c.q.d.

Ahora podemos dar el concepto de curva unicursal:

DEFINICIÓN XIV.5.4

*Se llama **curva unicursal** a una clase dada de equivalencia de aplicaciones racionales propias, respecto de la relación de equivalencia de la definición XIV.5.3. A estas aplicaciones se les llama **representaciones paramétricas propias** de la curva.*

Una curva unicursal queda, pues, determinada al dar una aplicación racional propia. Sus representaciones paramétricas propias se deducen de esta representación particular por medio de homografías no degeneradas cualesquiera.

Se plantea una última cuestión. Llamemos *imagen* de una aplicación racional cualquiera, al conjunto de puntos correspondientes a todos los valores del parámetro, para esta aplicación. (La imagen es, pues, una parte de \mathbf{C}^2 .) ¿Puede ser esta imagen igual a la imagen de una aplicación racional *propia*? ⁽²⁾. La respuesta a esta pregunta es positiva, y viene dada por un caso particular de un teorema general debido a Lüroth.

⁽¹⁾ Si tuviéramos $\varphi = 0$ (de donde $\psi \neq 0$) tendríamos $\Delta = \psi$, $\varphi_1 = 0$ y $\psi_1 = 1$: el razonamiento seguiría siendo válido. Asimismo si $\psi = 0$ (de donde $\varphi \neq 0$).

⁽²⁾ Se observará que, según XIV.5.2, dos aplicaciones racionales propias que tengan la misma imagen (salvo para un número finito de puntos) son equivalentes, lo que, a priori, no es evidente. Una parte de \mathbf{C}^2 es, pues, a lo sumo, la imagen de una curva unicursal.

Ejercicios

CAPÍTULO I

1. Sean E un conjunto y A, B dos partes de E , y se define la aplicación

$$f: \mathcal{P}(E) \rightarrow \mathcal{P}(A) \times \mathcal{P}(B) \quad \text{por} \quad f(X) = (X \cap A, X \cap B) \quad \text{para} \quad X \subset E.$$

¿Qué condición deben verificar A y B para que f sea inyectiva? ¿Para que f sea epiyectiva?

2. Sean A, B, C, D conjuntos, $f: A \rightarrow B, g: B \rightarrow C, h: C \rightarrow D$ aplicaciones. Si $g \circ f$ y $h \circ g$ son biyecciones, entonces f, g y h son biyecciones.

3. Sean A, B, C conjuntos, $f: A \rightarrow B, g: B \rightarrow C, h: C \rightarrow A$ aplicaciones,

$$\zeta = f \circ h \circ g, \quad \psi = g \circ f \circ h, \quad \varphi = h \circ g \circ f.$$

Si dos de las aplicaciones φ, ψ, ζ son inyectivas (resp. epiyectivas) y la tercera es epiyectiva (resp. inyectiva), entonces f, g y h son biyecciones.

4. Sean E y F dos conjuntos y f una aplicación de E en F . Se definen las aplicaciones $\hat{f}: \mathcal{P}(E) \rightarrow \mathcal{P}(F)$ y $\check{f}: \mathcal{P}(F) \rightarrow \mathcal{P}(E)$ por medio de las fórmulas

$$\hat{f}(A) = f(A) \quad (A \subset E) \quad \text{y} \quad \check{f}(B) = f^{-1}(B) \quad (B \subset F).$$

Demostrar que \hat{f} es inyectiva (resp. epiyectiva) si, y sólo si, f es inyectiva (resp. epiyectiva). ¿Qué condición debe verificar f para que \check{f} sea inyectiva (resp. epiyectiva)?

- *5. Sean X_1, X_2, \dots, X_n n conjuntos. Para toda parte H de \mathbf{N}_n^* , se pone

$$P_H = \bigcup_{i \in H} X_i \quad \text{y} \quad Q_H = \bigcap_{i \in H} X_i.$$

Sea \mathcal{E}_k el conjunto de las partes de \mathbf{N}_n^* que poseen k elementos ($0 \leq k \leq n$). Probar que se verifica:

$$a) \text{ si } k \leq \frac{n+1}{2}, \quad \bigcup_{H \in \mathcal{E}_k} Q_H \supset \bigcap_{H \in \mathcal{E}_k} P_H; \quad b) \text{ si } k \geq \frac{n+1}{2}, \quad \bigcup_{H \in \mathcal{E}_k} Q_H \subset \bigcap_{H \in \mathcal{E}_k} P_H.$$

6. Sea \mathcal{A} un conjunto de aplicaciones de un conjunto E en sí mismo, y \mathcal{F} la parte de $\mathcal{P}(E)$ formada por los conjuntos $X \subset E$ tales que $f(X) \subset X$ para toda $f \in \mathcal{A}$. Probar que, para la relación de inclusión, toda parte de \mathcal{F} admite un supremo y un ínfimo en \mathcal{F} .

7. Sea E un conjunto, y Ω el conjunto de las relaciones de orden definidas en E (se comprueba que Ω es también un conjunto). Se dice que el orden Γ_2 es *menos fino* que el orden Γ_1 y se escribe $\Gamma_2 \leq \Gamma_1$, si el grafo de Γ_1 contiene al grafo de Γ_2 . Probar que \leq es una relación de orden en Ω , y caracterizar los elementos *maximales* del conjunto Ω así ordenado.

(Se recuerda que el grafo de la relación de orden \leq definida en E es el conjunto de los $(x, y) \in E \times E$ tales que $x \leq y$.)

*8. Sea E un conjunto, eventualmente vacío, cuyos elementos son conjuntos. Se dice que E es *transitivo* si la relación $x \in E$ implica $x \subset E$.

a) Si X designa un conjunto cualquiera, se define $\mathcal{P}_n(X)$ por $\mathcal{P}_0(X) = X$ y

$$\mathcal{P}_{n+1}(X) = \mathcal{P}[\mathcal{P}_n(X)]$$

para todo $n \in \mathbf{N}$. Probar que, si X es transitivo, el conjunto $E = \bigcup_{n \in \mathbf{N}} \mathcal{P}_n(X)$ es transitivo. Caso en que $X = \emptyset$.

b) Si E es transitivo, demostrar que $E \cup \{E\}$ es transitivo. Si $(E_i)_{i \in I}$ es una familia de conjuntos transitivos, los conjuntos $\bigcup_{i \in I} E_i$ y $\bigcap_{i \in I} E_i$ son transitivos.

9. Para que un conjunto E sea finito es necesario y suficiente que toda parte no vacía de $\mathcal{P}(E)$ posea un elemento maximal respecto de la relación de inclusión.

(Si E no es finito, se utilizará una inyección: $\mathbf{N} \rightarrow E$.)

10. Calcular el número λ_n de regiones del plano \mathbf{R}^2 determinadas por n rectas tales que nunca tres de ellas sean concurrentes, y que nunca dos de ellas sean paralelas.

Calcular asimismo el número μ_n de regiones del espacio \mathbf{R}^3 determinadas por n planos tales que no haya dos de ellos paralelos, ni tres de ellos que contengan una recta en común, ni cuatro de entre ellos que contengan un punto en común.

11. Consideremos n puntos cualesquiera del espacio \mathbf{R}^3 . Los planos que pasan por tres de tales puntos se cortan según v_n rectas. Calcular v_n en función de n .

12. Consideremos n puntos cualesquiera del espacio \mathbf{R}^3 . Las esferas que pasan por 4 de estos puntos se cortan según v_n circunferencias (reales o imaginarias) y según μ_n puntos. Calcular v_n y μ_n en función de n .

13. Se da $f(x) = \sum_{k=0}^m \frac{x^{k+1}}{k+1} \binom{m}{k}$. Calcular $f'(x)$ y deducir el valor de $f(1)$.

14. Se da $f(x) = \sum_{k=1}^n (-1)^{k-1} \frac{1}{k} \binom{n}{k} x^k$. Calcular $xf'(x)$ y deducir el valor de $f(1)$.

15. Si suponemos $p \leq n$, $q < p$ y $p \geq 1$, probar que

$$\sum_{k=q+1}^{n-p+q+1} \binom{n-k}{p-q-1} \binom{k-1}{q} = \binom{n}{p}.$$

Método. Sea A el conjunto de las partes de \mathbf{N}_n^* con p elementos. Para $X \in A$ escribamos $X = \{i_1, i_2, \dots, i_p\}$ con $1 \leq i_1 < i_2 < \dots < i_p \leq n$. Para todo k tal que

$$q+1 \leq k \leq n-p+q+1,$$

sea A_k el conjunto de los $X \in A$, $X = \{i_1, i_2, \dots, i_p\}$, tales que $i_{q+1} = k$. Calcular $\text{card}(A_k)$ y probar que los A_k constituyen una partición de A .

16. Sea E un conjunto finito con np elementos. Calcular el número de particiones de E en n partes con p elementos cada una.

*17. Sea E un conjunto finito con n elementos, \mathcal{F} el conjunto de las aplicaciones de $\mathcal{P}(E)$ en \mathbf{R} . Para toda parte A de E , se establece $|A| = \text{card}(A)$. Se definen las aplicaciones φ y ψ de \mathcal{F} en \mathcal{F} por medio de

$$[\varphi(f)](A) = \sum_{B \subset A} f(B), \quad [\psi(g)](A) = \sum_{B \subset A} (-1)^{|A \setminus B|} g(B) \quad \text{para } f, g \in \mathcal{F}.$$

Probar que φ y ψ son biyecciones recíprocas una de otra. (Se estudiarán las aplicaciones $\varphi \circ \psi$ y $\psi \circ \varphi$.)

18. Sea E un conjunto finito con n elementos. Un *recubrimiento de E con p elementos* es una aplicación $k \mapsto X_k$ de \mathbf{N}_p^* en $\mathcal{P}(E) \setminus \{\emptyset\}$ tal que $\bigcup_{1 \leq k \leq p} X_k = E$. Sea $R_{n,p}$ el número de recubrimientos de E con p elementos. Calcular $R_{n,2}$ y $R_{n,3}$.

19. Para todo par de enteros $n \geq 1$ y $\alpha \geq 0$, se pone: $S_{n,\alpha} = \sum_{p=1}^n p^\alpha$. Aplicar la fórmula del binomio para calcular $S_{n+1,\alpha+1}$ y deducir la relación de recurrencia

$$(n+1)^{\alpha+1} - 1 = \sum_{\lambda=0}^{\alpha} \binom{\alpha+1}{\lambda} S_{n,\lambda}.$$

Calcular $S_{n,0}$, $S_{n,1}$, $S_{n,2}$ y $S_{n,3}$.

20. Número de soluciones $(x, y, z) \in \mathbf{N}^3$ de la ecuación $x + 2y + 5z = n$ ($n \in \mathbf{N}$).

(Número de formas de pagar 10 francos franceses por medio de monedas de 1, 2 y 5 céntimos de franco.)

(Ecole polytechnique.)

[Se observará que el número u_n que se busca verifica la igualdad

$$\frac{1}{(1-x)(1-x^2)(1-x^5)} = \sum_{n=0}^{\infty} u_n x^n;$$

y se descompondrá esta fracción racional en elementos simples en \mathbf{C} a fin de obtener su desarrollo en serie formal (cf. Cap. VII). Generalización evidente.]

21. Fijado de antemano el entero $p \in \mathbf{N}$, calcular

$$\sum_{x+y=p(x,y \in \mathbf{N})} x^2 y^2 \quad \text{y} \quad \sum_{x+y+z=p(x,y,z \in \mathbf{N})} xyz.$$

22. Sea $S_{n,p}$ el número de aplicaciones epiyectivas de \mathbf{N}_n^ en \mathbf{N}_p^* ($n \geq p > 1$).

a) Establecer la relación

$$(1) \quad p^n = S_{n,p} + \binom{p}{1} S_{n,p-1} + \cdots + \binom{p}{p-1}.$$

(Se evaluará de dos maneras diferentes el número de aplicaciones de \mathbf{N}_n^* en \mathbf{N}_p^* .)

b) Establecer las relaciones:

$$(2) \quad \begin{cases} 2^p \binom{n}{p} = \binom{n}{0} \binom{n}{p} + \binom{n}{1} \binom{n-1}{p-1} + \cdots + \binom{n}{p} \binom{n-p}{0} \\ 0 = \binom{n}{0} \binom{n}{p} - \binom{n}{1} \binom{n-1}{p-1} + \cdots + (-1)^p \binom{n}{p} \binom{n-p}{0} \end{cases} \quad (p \leq n, n \geq 1)$$

(Se podrá realizar el cálculo directo o bien desarrollar de dos maneras

$$(1 + X + X)^n \quad \text{y} \quad (1 + X - X)^n.)$$

c) Deducir la fórmula

$$(3) \quad S_{n,p} = p^n - \binom{p}{1} (p-1)^n + \binom{p}{2} (p-2)^n + \cdots + (-1)^{p-1} \binom{p}{p-1}.$$

(Se sumarán las relaciones obtenidas reemplazando sucesivamente p por $p-1, p-2, \dots, 1$ en (1); se podrá utilizar también el ejercicio IX.8.)

d) Para $n \geq p \geq 2$, probar que

$$S_{n,p} = p(S_{n-1,p} + S_{n-1,p-1}).$$

Deducir

$$S_{n+1,n} = \frac{n}{2} (n+1)! \quad \text{y} \quad S_{n+2,n} = \frac{n(3n+1)}{24} (n+2)!$$

e) Hallar el número $P_{n,p}$ de relaciones de equivalencia \mathcal{R} que admite un conjunto finito E con n elementos, tales que $\text{card}(E/\mathcal{R}) = p$ ($1 \leq p \leq n$).

23. Sea p_n el número de permutaciones u del conjunto \mathbf{N}_n^ tales que, para todo $x \in \mathbf{N}_n^*$, verifican $u(x) \neq x$.

a) Establecer la relación:

$$(1) \quad n! = p_n + \binom{n}{1} p_{n-1} + \cdots + \binom{n}{n-2} p_2 + 1.$$

(Se calculará de dos maneras el número de permutaciones de \mathbf{N}_n^* .)

b) Procediendo como en el ejercicio anterior, o bien utilizando el ejercicio IX.8, deducir

$$p_n = n! - \binom{n}{1} (n-1)! + \binom{n}{2} (n-2)! + \cdots + (-1)^n.$$

c) Probar que

$$\lim_{n \rightarrow \infty} \frac{ep_n}{n!} = 1.$$

(e designa el número real $\sum_{p=0}^{\infty} \frac{1}{p!}$).

24. Sea $q_{n,k}$ el número de las aplicaciones $u: \mathbf{N}_k^* \rightarrow \mathbf{N}_n^*$, estrictamente crecientes, que transforman todo número par en un número par y todo número impar en un número impar. Probar que

$$(1) \quad q_{n,k} = q_{n-1,k-1} + q_{n-2,k},$$

y deducir

$$(2) \quad q_{n,k} = \binom{\left\lceil \frac{n+k}{2} \right\rceil}{k},$$

(en donde $\left[\frac{n+k}{2} \right]$ designa la *parte entera* de $\frac{n+k}{2}$).

25. Sean E y F dos conjuntos finitos, y se ordenan por inclusión los conjuntos $\mathcal{P}(E)$ y $\mathcal{P}(F)$. Calcular, en función de $m = \text{card}(E)$ y de $n = \text{card}(F)$, el número de las aplicaciones *crecientes* de $\mathcal{P}(E)$ en $\mathcal{P}(F)$.

(Indicación: Se calculará el número de aplicaciones no crecientes.)

CAPÍTULO II

1. Sean G' y G'' dos subgrupos del grupo G . Si $G' \cup G'' = G$, se tiene $G' = G$ o $G'' = G$.
2. Determinar todos los grupos de 4 ó 6 elementos. ¿Cuáles son conmutativos?
3. Sea H un subgrupo del grupo G . Si $[G : H] = 2$, H es normal.
4. Si H y K son dos subgrupos finitos del grupo G , cuya intersección se reduce a $\{e\}$, probar que $\text{card}(H.K) = \text{card}(H) \cdot \text{card}(K)$. (Se recuerda que

$$H.K = \{ h.k \}_{h \in H, k \in K} .)$$

[Atención : $H.K$ no es forzosamente un subgrupo y se podrán hallar ejemplos en los que $H.K$ engendra un subgrupo infinito de G .]

5. Sea G un grupo de orden pq , en donde p es primo y verifica $p > q$. Probar, utilizando el ejercicio anterior, que G tiene a lo sumo un subgrupo de orden p y deducir que (si existe) dicho subgrupo es normal.

6. Sea E un conjunto provisto de una ley interna asociativa. Si E es finito, y si todos los elementos de E son regulares, E es un grupo.

7. Sea E un conjunto provisto de una ley interna asociativa. Si existe un elemento e de E tal que, para todo $x \in E$, verifica $xe = x$, y si, para todo $x \in E$, existe un $y \in E$ que verifica $xy = e$, E es un grupo.

8. Estudiar la ley de composición interna $a \top b = a + b - ab$ sobre \mathbf{Q} , \mathbf{R} , o \mathbf{C}

(Ecole polytechnique.)

9. Estudiar la ley $(x, y) \mapsto \frac{x+y}{1+(xy/c^2)}$ en el intervalo $I =]-c, +c[$ de \mathbf{R} . Probar que define una estructura de grupo conmutativo.

(Ecole polytechnique.)

(Se observará que $x \mapsto \arg t \ x/c$ es un isomorfismo de este grupo en el grupo aditivo de los números reales; se estudiará asimismo la ley $(x, y) \mapsto x + y/1 - (xy/c^2)$).

10. Sea A un entero > 0 , no cuadrado perfecto. Se designa por \mathcal{F} el conjunto de los pares de enteros x, y tales que $x^2 - Ay^2 = 1$ y que $y \neq 0$. Se supone que \mathcal{F} es no vacío ⁽¹⁾ y se dota a \mathcal{F} de la ley interna \top definida por

$$(X, Y) = (x, y) \top (x', y') \quad \text{si} \quad (x + y\sqrt{A})(x' + y'\sqrt{A}) = X + Y\sqrt{A} .$$

a) Probar que \top es una ley de grupo abeliano en \mathcal{F}

b) Sean (x, y) y (x_0, y_0) dos elementos de \mathcal{F} tales que $x > 0, y > 0, x_0 > 0, y_0 > 0$. Se define $(x, y) \top (x_0, -y_0) = (x_1, y_1)$. Probar que $x_0 < x$ implica $x_1 < x$.

c) Se designa por \mathcal{F}^+ al conjunto de los $(x, y) \in \mathcal{F}$ tales que $x > 0$ e $y > 0$. Probar que existe un $(x_0, y_0) \in \mathcal{F}^+$ para el cual $x = x_0$ es mínimo y probar que este elemento es un generador del subgrupo \mathcal{F}^+ .

d) Dar un sistema de generadores de \mathcal{F} .

⁽¹⁾ La parte difícil del estudio de la ecuación diofántica $x^2 - Ay^2 = 1$ (llamada ecuación de Fermat) consiste en demostrar que $\mathcal{F} \setminus \{(1, 0), (-1, 0)\}$ es no vacío.

e) Hallar todas las soluciones en números enteros x, y de las ecuaciones

$$x^2 - 2y^2 = 1, \quad x^2 - 3y^2 = 1.$$

***11.** Sea G un grupo, indicado multiplicativamente, y sea $[G, G]$ el subgrupo de G engendrado por los elementos de la forma

$$x^{-1} y^{-1} xy \quad (x \in G, y \in G).$$

a) Probar que $[G, G]$ es normal en G y que $G/[G, G]$ es abeliano.

b) Probar que $[G, G]$ tiene la propiedad «universal» siguiente: Para todo grupo abeliano H , y todo homomorfismo $f: G \rightarrow H$, existe un homomorfismo y sólo uno $\bar{f}: G/[G, G] \rightarrow H$, tal que $\bar{f} \circ p = f$, en donde $p: G \rightarrow G/[G, G]$ es la aplicación canónica

$$\begin{array}{ccc} G & \xrightarrow{p} & G/[G, G] \\ & \searrow f & \downarrow \bar{f} \\ & & H \end{array}$$

(Esta propiedad significa que $[G, G]$ es el menor subgrupo normal K de G tal que G/K es abeliano). A $[G, G]$ se le llama *grupo de los conmutadores de G* .

c) Hallar el grupo $[G, G]$ cuando G es uno de los grupos siguientes: $\mathfrak{S}_3, \mathfrak{S}_4, \mathcal{A}_4, \mathcal{A}_5$.

12. Sea G un grupo, A el grupo de todos sus automorfismos, Γ [resp. Δ] el grupo de las traslaciones por la izquierda (resp. por la derecha) de G ; se observa que A, Γ y Δ son subgrupos del grupo \mathfrak{S}_G de las permutaciones de G .

a) Probar que $A\Gamma = \Gamma A$, y que $\Omega = A\Gamma$ es un subgrupo de \mathfrak{S}_G .

b) Probar que Γ es normal en Ω , y que todo automorfismo de Γ es de la forma $\gamma \mapsto \sigma \gamma \sigma^{-1}$, en donde $\sigma \in A$.

c) Probar que Δ es normal en Ω , y que $\Gamma \cap \Delta$ es el centro de cada uno de los grupos Γ, Δ .

Indicación práctica: Será cómodo designar por Γ_g la traslación por la izquierda $x \mapsto g.x$, por Δ_h la traslación por la derecha, $x \mapsto xh$, y por A_h el automorfismo interno $x \mapsto h x h^{-1}$.

13. En lo que sigue, si $A = \{a_1, \dots, a_p\}$ es una parte de \mathbf{N}_n^* , designaremos por $[a_1, \dots, a_p]$ el ciclo s que opera sobre A tal que:

$$s(a_1) = a_2, \dots, s(a_{p-1}) = a_p, \quad s(a_p) = a_1.$$

a) Probar que, para toda permutación $\sigma \in \mathfrak{S}_n$, se tiene:

$$\sigma[a_1 \dots a_p] \sigma^{-1} = [\sigma(a_1) \sigma(a_2) \dots \sigma(a_p)].$$

b) Probar que \mathfrak{S}_n está engendrado por las $n-1$ trasposiciones

$$(12), (13), \dots, (1n),$$

y por las $n-1$ trasposiciones

$$(12), (23), \dots, (n-1, n).$$

c) Probar que \mathfrak{S}_n está engendrado por las dos permutaciones

$$(12) \quad y \quad (1\ 2\ 3\ \dots\ n).$$

d) Probar que \mathcal{A}_n está engendrado por los ciclos de orden 3. (Probarlo en primer lugar para los productos de *dos* trasposiciones.)

14. Probar que el grupo alternado \mathcal{A}_n ($n \geq 3$) es $n-2$ veces transitivo sobre \mathbf{N}_n^* .

15. Si H y K son dos subgrupos del grupo G , probar que el subconjunto $H.K$ de G contiene exactamente $[H] \times [K]/[H \cap K]$ elementos. (Definir una relación de equivalencia en $H.K$.)

16. Sea G un grupo infinito que admita un subgrupo H de índice finito, y sea

$$K = \bigcap_{g \in G} (gHg^{-1})$$

la intersección de los conjugados de H . Probar que existe una parte finita F de G tal que $K = \bigcap_{g \in F} (gHg^{-1})$; probar que K es normal en G y que el índice $[G : K]$ es finito.

17. Se dice que el grupo G verifica la condición (I) si, para todo $x \in G$, se tiene $x^2 = e$.

a) Probar que tales grupos son abelianos (utilizar la relación $x^2y^2 = xyxy$).

b) Sea (G_i) ($1 \leq i \leq n$) una familia finita de grupos que verifiquen la condición (I). Probar que su producto $G = \prod G_i$ también verifica la condición (I).

c) Partiendo del grupo G con dos elementos, probar que para todo $n \in \mathbf{N}$, se puede construir un grupo G_n de orden 2^n que verifique (I).

Interpretar G_n como un grupo de transformaciones, estudiando todas las simetrías asociadas a una referencia ortonormal de E^n (se empezará por los casos $n = 2$, $n = 3$).

18. a) Sea $\text{SO}(3)$ el grupo de las rotaciones de \mathbf{R}^3 que dejan fijo el origen. Probar que las rotaciones, con ejes que pasan por O , y cambian dos puntos diametralmente opuestos de la esfera unidad S^2 de \mathbf{R}^3 , son elementos de orden 2 de $\text{SO}(3)$.

b) Se considera un subgrupo G de $\text{SO}(3)$ que sea *transitivo* sobre S^2 . Probar que G contiene un elemento de orden 2, y deducir de ello que contiene *todos* los elementos de orden 2 y, por lo tanto, que es idéntico a $\text{SO}(3)$.

*19. Un grupo G es *simple* si no admite más subgrupos normales que $\{e\}$ y G .

a) Probar que el grupo alternado \mathcal{A}_5 es simple (utilizar el ejercicio 14).

b) Probar que el grupo $\text{SO}(3)$ es simple (utilizar el ejercicio 18).

20. Se llama *razón doble* (o razón anarmónica) de cuatro números complejos distintos x, y, z, t al número $B(x, y, z, t) = \frac{z-x}{z-y} : \frac{t-x}{t-y}$.

a) Demostrar que $(x, y, z, t) \mapsto B(x, y, z, t)$ es una aplicación de una parte de \mathbf{C}^4 , que se precisará, en el conjunto $E = \mathbf{C}^* - \{1\}$.

b) A cada permutación $\sigma \in \mathfrak{S}_4$ se le asocia el número $B[\sigma(x), \sigma(y), \sigma(z), \sigma(t)]$. Probar que este número depende únicamente de σ y del número $\lambda = B(x, y, z, t)$. Se escribirá

$$B[\sigma(x), \sigma(y), \sigma(z), \sigma(t)] = \varphi(\sigma, \lambda).$$

c) Hacemos operar \mathfrak{S}_4 sobre $\mathbf{C}^* - \{1\}$ por medio de la ley de composición $(\sigma, \lambda) \mapsto \varphi(\sigma, \lambda)$. Discutir, según los distintos valores de λ , el número de elementos de su órbita y precisar su grupo de isotropía (hay que distinguir tres casos).

d) Para que cuatro números complejos distintos a, b, c, d sean los afijos de cuatro puntos alineados o cocíclicos (e.d. sobre una misma circunferencia) es necesario y suficiente que su razón doble sea real. Interpretar entonces el signo de dicha razón doble.

21. Se designa por $\tilde{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$ al conjunto obtenido completando \mathbf{C} con un punto del «infinito»; al conjunto $\tilde{\mathbf{C}}$ se le llama «esfera de Riemann». Toda función homográfica

$$f: z \mapsto \frac{az + b}{cz + d} \quad (ad - bc \neq 0)$$

se prolonga, por convenio, a los puntos $-d/c$ e ∞ de $\tilde{\mathbf{C}}$ con la ayuda de las fórmulas

$$(1) \quad f\left(-\frac{d}{c}\right) = \infty \quad f(\infty) = \frac{a}{c} \quad \text{cuando} \quad c \neq 0, \quad \text{y} \quad f(\infty) = \infty \quad \text{si} \quad c = 0.$$

a) Probar que el conjunto de las aplicaciones de la forma

$$z \mapsto \frac{az + b}{cz + d}, \quad \text{con} \quad ad - bc \neq 0,$$

constituye un grupo G de biyecciones de $\tilde{\mathbf{C}}$ (a este grupo se le llama *grupo circular*).

¿Cuál es el subgrupo G_0 que deja invariante el origen? ¿Y el subgrupo G_∞ que deja invariante el punto del infinito?

b) La noción de razón doble (ver ejercicio II.20) se extiende a $\tilde{\mathbf{C}}$ utilizando los convenios (1):

Probar que una aplicación $f: \tilde{\mathbf{C}} \rightarrow \tilde{\mathbf{C}}$ es una homografía si, y sólo si, se verifica $B(f(x), f(y), f(z), f(t)) = B(x, y, z, t)$ cualesquiera que sean los puntos distintos x, y, z, t de $\tilde{\mathbf{C}}$.

22. Sea \mathcal{H} el grupo de las homografías

$$z \mapsto \frac{az + b}{cz + d} \quad (a, b, c, d \in \mathbf{C}, ad - bc \neq 0)$$

de la esfera de Riemann $\tilde{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$.

a) Sea $H(z_0, z_1)$ el subgrupo de \mathcal{H} formado por las homografías h que dejan fijos dos puntos dados y distintos z_0, z_1 de $\tilde{\mathbf{C}}$. Probar que $H(z_0, z_1)$ es isomorfo al grupo multiplicativo $\mathbf{C}^* = \mathbf{C} - \{0\}$ y deducir de este resultado que todo subgrupo finito de $H(z_0, z_1)$ es cíclico. (Se estudiará el transformado de $H(z_0, z_1)$ por medio del automorfismo interno $h \mapsto \varphi h \varphi^{-1}$, en donde φ designa la homografía definida por

$$\varphi(z) = \frac{z - z_0}{z - z_1}.$$

b) Probar que las homografías que admiten como *único* punto fijo un punto dado z_0 de $\tilde{\mathbf{C}}$ constituyen un subgrupo de \mathcal{H} , a saber $G(z_0)$, isomorfo al grupo aditivo \mathbf{C} . Deducir que $G(z_0)$ no admite más subgrupo finito que el reducido a la identidad.

(Se estudiará el transformado de $G(z_0)$ por el automorfismo interno $h \mapsto \psi h \psi^{-1}$, en donde ψ designa la homografía definida por

$$\psi(z) = \frac{1}{z - z_0}.$$

23. Sea \mathcal{M} el conjunto de las homografías de la esfera de Riemann $\tilde{\mathbf{C}}$, de la forma

$$z \mapsto \frac{az + b}{cz + d}, \quad \text{con} \quad ad - bc = 1, \quad a, b, c, d \in \mathbf{Z}.$$

a) Probar que \mathcal{M} es un grupo.

b) Para todo $z_0 \in \mathbf{C}$, se designa por Ω_{z_0} a la órbita de z_0 por \mathcal{M} . Probar que si $\text{Im}(z_0) > 0$, z_0 es un punto aislado de Ω_{z_0} .

(Se calculará la parte imaginaria de $\frac{az+b}{cz+d}$ en función de la de z .)

(Ecole polytechnique.)

24. Sea $z \rightarrow \varphi(z) = \frac{az+b}{cz+d}$ una transformación circular de la esfera de Riemann $\tilde{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$.

El grupo G_φ engendrado por φ en el grupo circular, opera sobre $\tilde{\mathbf{C}}$ por medio de la ley

$$(\psi, z) \mapsto \psi(z) \quad (\psi \in G_\varphi, z \in \tilde{\mathbf{C}}).$$

Probar que si existe una órbita finita de $\tilde{\mathbf{C}}$ con dos elementos, por lo menos, φ es de orden finito (es decir G_φ es un grupo finito).

25. Se hace operar un grupo G sobre sí mismo por medio de los automorfismos internos de G . Escribir la ecuación de las clases. Caracterizar las órbitas reducidas a un elemento.

Deducir que si $[G] = p^n$, en donde p es primo, el centro de G no está reducido al elemento neutro.

26. Sea K un cubo de \mathbf{R}^3 , de centro O . Se designan por $\alpha, \beta, \gamma, \delta$ las cuatro diagonales de K , por \mathcal{D}_K el grupo de las isometrías de \mathbf{R}^3 que dejan a K globalmente invariante, y por \mathcal{S}_K al grupo de los desplazamientos de \mathbf{R}^3 que dejan a K globalmente invariante.

a) Haciendo operar \mathcal{D}_K sobre el conjunto $\Delta = \{\alpha, \beta, \gamma, \delta\}$, probar que \mathcal{D}_K posee 48 elementos.

b) Probar que $\mathcal{S}_K \neq \mathcal{D}_K$. Haciendo operar \mathcal{S}_K sobre Δ , probar que \mathcal{S}_K es isomorfo al grupo de las permutaciones de Δ .

c) Probar que con los vértices de K se pueden formar dos tetraedros regulares T_1, T_2 . Determinar el subgrupo de los $\sigma \in \mathcal{D}_K$ tales que $\sigma(T_i) = T_i$ ($i = 1, 2$) y el subgrupo de $\sigma \in \mathcal{S}_K$ tales que $\sigma(T_i) = T_i$. ¿Cuáles son los elementos s de \mathcal{D}_K tales que $s(T_1) = T_2$ (y $s(T_2) = T_1$)?

d) ¿Cuáles son los subgrupos de \mathcal{D}_K isomorfos al grupo de las isometrías planas que dejan globalmente invariante un cuadro? Estos subgrupos, ¿se hallan contenidos en \mathcal{S}_K ? ¿Cuáles son los subgrupos de \mathcal{D}_K isomorfos al grupo de los desplazamientos que dejan globalmente invariante un cuadro?

27. Para cada entero $n \geq 1$ se identifica \mathbf{R}^n con el hiperplano $x_{n+1} = 0$ de \mathbf{R}^{n+1} ; y se define por recurrencia un simplex regular de \mathbf{R}^n , designado por $T_n = \{a_0, a_1, \dots, a_n\}$, de la forma siguiente: T_1 está formado por dos puntos distintos a_0, a_1 de \mathbf{R} ; y si

$$T_{n-1} = \{a_0, a_1, \dots, a_{n-1}\}$$

está definido, tal que $T_{n-1} \subset \mathbf{R}^{n-1}$, se establece $T_n = T_{n-1} \cup \{a_n\}$, en donde a_n designa un punto de \mathbf{R}^n tal que, para todo $i = 1, 2, \dots, n-1$, se verifique $d(a_n, a_i) = d(a_0, a_1)$ (en donde d designa la distancia euclídea).

a) Probar que la construcción de T_n es posible, que los puntos a_0, a_1, \dots, a_n son afínmente libres en \mathbf{R}^n (cf. § VIII.6) y que el isobaricentro Ω_n de los puntos de T_n equidista de todos los puntos de T_n .

b) Probar que el grupo de las isometrías de \mathbf{R}^n que dejan a T_n globalmente invariante es isomorfo a \mathfrak{S}_n , y que el grupo de los desplazamientos que dejan a T_n globalmente invariante es isomorfo a \mathcal{A}_n .

CAPÍTULO III

1. Todo anillo íntegro finito es un cuerpo.

2. Sea A un anillo unífero no necesariamente, conmutativo. Se llama *ideal por la izquierda* de A (resp. *ideal por la derecha* de A) a todo subgrupo aditivo de A , por ejemplo \mathfrak{b} , tal que

$$\forall \lambda \in A \quad \forall x \in \mathfrak{b}, \quad \lambda x \in \mathfrak{b} \quad (\text{resp. } x\lambda \in \mathfrak{b}).$$

Un ideal *bilátero* es un ideal que lo es simultáneamente por la derecha y por la izquierda.

a) Probar que es posible definir el anillo cociente de A por un ideal bilátero, y que se obtiene un teorema análogo a III.4.4.

b) Si \mathfrak{b} es un ideal bilátero de A y si el anillo cociente A/\mathfrak{b} es un cuerpo, ¿en qué conjunto de ideales (ordenado por inclusión) \mathfrak{b} es maximal?

3. Sea A un anillo sin divisores de 0, tal que todo subgrupo aditivo de A sea un ideal por la izquierda de A . Probar que A es isomorfo a \mathbf{Z} (considerar los grupos engendrados por un elemento de A).

4. Los números de la forma $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ ($a, b, c, d \in \mathbf{Q}$) constituyen un subcuerpo de \mathbf{R} .

5. Sea p un número primo ≥ 3 tal que $p \equiv 3 \pmod{4}$. Probar que en el cuerpo $F_p = \mathbf{Z}/p\mathbf{Z}$, la relación $a^2 + b^2 = 0$ implica $a = b = 0$ (cf. ejemplo 3 del § IV.6, *criterio de Euler*). Generalizando la construcción de \mathbf{C} a partir de \mathbf{R} , deducir la existencia de un cuerpo (designado por F_{p^2}) tal que $F_p \subset F_{p^2}$ y que $-\bar{1}$ posea una raíz cuadrada en F_{p^2} .

*6. Sea K un cuerpo y $\sigma : K \rightarrow K$ un isomorfismo de K en sí mismo. Sea \mathcal{G} el conjunto de las sucesiones $\alpha = (\alpha_n)_{(n \in \mathbf{Z})}$ de elementos de K tales que existe un $n_0 \in \mathbf{Z}$ que verifica $\alpha_n = 0$ para $n < n_0$. Se dota a \mathcal{G} de la estructura de grupo aditivo inducida por el grupo producto $K^{\mathbf{Z}}$, y de la multiplicación siguiente:

$$\alpha \cdot \beta = \gamma, \quad \text{con} \quad \gamma_n = \sum_{p+q=n} \alpha_p \sigma^p(\beta_q).$$

Probar que \mathcal{G} es un cuerpo. ¿Es conmutativo?

K designa el cuerpo F_{p^2} del ejercicio 5). Hallar un automorfismo conveniente de K que permita deducir la existencia de un cuerpo no conmutativo de característica p , para $p \equiv 3 \pmod{4}$.

Si $p \equiv 1 \pmod{4}$, probar que existe un cuerpo F_{p^2} , conmutativo, de cardinal p^2 que contiene a F_p .

(Para terminar este ejercicio, probar que, para todo número primo $p \geq 3$, existe un cuerpo no conmutativo de característica p .)

7. Sea \mathcal{D} un ideal del anillo \mathcal{C} de las funciones continuas: $[0, 1] \rightarrow \mathbf{R}$. Se supone que, para cada punto $x \in [0, 1]$, existe una $f \in \mathcal{C}$ tal que $f(x) \neq 0$.

Probar que existe una parte *finita* de \mathcal{D} que posee la misma propiedad y deducir la existencia de una función $f \in \mathcal{D}$ que no se anula en ningún punto de $[0, 1]$; deducir que $\mathcal{D} = \mathcal{C}$ (utilizar la compacidad de $[0, 1]$). ¿Cuáles son los ideales maximales de \mathcal{C} ? Buscar otros ideales de \mathcal{C} .

8. Sean A, B dos anillos conmutativos, uníferos, $h : A \rightarrow B$ un homomorfismo, y J un ideal de B . Probar que, si h es epiyectiva y J es maximal, el ideal $h^{-1}(J)$ es maximal.

9. Probar que en un anillo conmutativo unífero A , el conjunto \mathcal{N} de los elementos nilpotentes es un ideal ($a \in A$ es nilpotente si existe un $n \in \mathbf{N}^*$ tal que $a^n = 0$). ¿Qué podemos decir del anillo cociente A/\mathcal{N} ? Deducir que, para todo ideal \mathcal{D} de A , el conjunto $v(\mathcal{D})$ de los $x \in A$ tales que existe un $n \in \mathbf{N}^*$ que verifica $x^n \in \mathcal{D}$, es un ideal de A .

10. Probar que el número de elementos de un cuerpo finito es una potencia de su característica.

11. Se designa por $G(n)$ al grupo de las unidades del anillo $\mathbf{Z}/n\mathbf{Z}$ ($n \in \mathbf{N}^*$).

a) Probar que $G(n_1 \cdot n_2 \dots n_k)$ es cíclico si, y sólo si, se satisfacen las siguientes condiciones: cada $G(n_i)$ es cíclico y los números (n_i) son, dos a dos, primos entre sí.

b) Si p es un número primo impar, $G(p)$ es cíclico (cf. § IV.5). Se designa por \bar{x} la clase, en $\mathbf{Z}/p\mathbf{Z}$, del entero $x \in \mathbf{Z}$. Establecer que, para todo generador $\bar{\zeta}$ de $G(p)$ (en donde $\zeta \in \mathbf{Z}$), existe un

entero $k \in \mathbf{Z}$ tal que $\bar{\theta} = \overline{\zeta + kp}$ es un generador de $G(p)$, y que $(\zeta + kp)^{p-1} - 1$ no es divisible por p^2 . Probar entonces que, para todo entero $\alpha \geq 1$, la clase de θ en $\mathbf{Z}/p^\alpha \mathbf{Z}$ es un generador de $G(p^\alpha)$.

c) Probar que $G(2^m)$ es cíclico si, y sólo si, $m = 1$ o $m = 2$. A este efecto se establecerá la relación $a^{2^m-2} \equiv 1 \pmod{2^m}$, válida para $m > 2$ y para a entero e impar.

d) De lo anterior deducir que $G(n)$ es cíclico si, y sólo si, n tiene una de las formas siguientes: $n = 2$, $n = 4$, $n = p^\alpha$, $n = 2p^\alpha$ (p primo e impar, $\alpha \geq 1$).

12. Sea E un conjunto y A el anillo $\mathbf{Z}/2\mathbf{Z}$. Para toda parte X de E , se define la aplicación $\chi_X : E \rightarrow A$ por

$$\chi_X(x) = \bar{0} \quad \text{si} \quad x \notin X \quad \text{y} \quad \chi_X(x) = \bar{1} \quad \text{si} \quad x \in X.$$

a) $X \mapsto \chi_X$ es una biyección de $\mathcal{P}(E)$ en el anillo $\mathcal{F}(E, A)$. Con la ayuda de esta biyección, se transporta sobre $\mathcal{P}(E)$ la estructura de anillo de $\mathcal{F}(E, A)$. ¿Qué operaciones corresponden, en $\mathcal{P}(E)$, a la suma y al producto de $\mathcal{F}(E, A)$?

b) Si E es finito, todo ideal de $\mathcal{F}(E, A)$ es de la forma Z_X , en donde $X \subset E$ y en donde

$$Z_X = \{ f \in \mathcal{F}(E, A) \mid \forall x \in X, f(x) = \bar{0} \}.$$

Deducir, en este caso, los ideales maximales.

c) Se considera ahora un anillo conmutativo unífero \mathcal{B} , tal que $x + x = 0$ y $x^2 = x$ para todo $x \in \mathcal{B}$. Probar que, si \mathcal{B} es íntegro, \mathcal{B} es isomorfo a $\mathbf{Z}/2\mathbf{Z}$. Deducir, en este caso, los ideales \mathfrak{p} de \mathcal{B} para los que el anillo \mathcal{B}/\mathfrak{p} es íntegro (ideales primos de \mathcal{B}).

*d) Se supone que el anillo \mathcal{B} verifica las hipótesis de c) y es finito. Sea \mathcal{M} el conjunto de los ideales maximales de \mathcal{B} . Se define la aplicación

$$\varphi : \mathcal{B} \rightarrow \mathcal{F}(\mathcal{M}, A)$$

por $\varphi(x) = \tilde{x}$, con $\tilde{x}(\mathfrak{p}) = \bar{0}$ si $x \in \mathfrak{p}$, $\tilde{x}(\mathfrak{p}) = \bar{1}$ si $x \notin \mathfrak{p}$ ($\mathfrak{p} \in \mathcal{M}$).

Probar que φ es un isomorfismo de anillos.

13. Sea T un conjunto y $\mathcal{A}(T)$ la \mathbf{C} -álgebra de las aplicaciones $f : T \rightarrow \mathbf{C}$ (el producto de las aplicaciones $x \mapsto g(x)$ y $x \mapsto f(x)$ es la aplicación $x \mapsto f(x)g(x)$). Se llama *carácter en $\mathcal{A}(T)$* a toda aplicación no nula $\chi : \mathcal{A}(T) \rightarrow \mathbf{C}$, que sea \mathbf{C} -lineal, y tal que $\chi(fg) = \chi(f)\chi(g)$ para $f, g \in \mathcal{A}(T)$.

1) Demostrar las propiedades siguientes para todo carácter χ .

- a) Si $(\forall x \in T, f(x) \neq 0)$, entonces $\chi(f) \neq 0$; b) $\chi(f) \in f(T)$; c) $(f \geq 0) \Rightarrow (\chi(f) \geq 0)$
d) $\chi(f) = \chi(f)$; e) $\chi(|f|) = |\chi(f)|$; f) $\chi(1) = I$, en donde I designa la aplicación constante igual a 1.

2) Si χ_1, \dots, χ_m son caracteres, distintos dos a dos, son linealmente independientes (es decir: las relaciones $\lambda_1, \dots, \lambda_m \in \mathbf{C}$ y $\sum_{i=1}^m \lambda_i \chi_i = 0$ implican las relaciones $\lambda_i = 0$ ($i = 1, 2, \dots, m$)).

*3) Se supone ahora que T es finito. Probar que, para todo carácter χ , existe un $x_0 \in T$ tal que, para todo $f \in \mathcal{A}(T)$, se verifica: $\chi(f) = f(x_0)$.

14. *Enteros de Gauss.* Sea A el anillo de los números complejos de la forma $a + ib$ ($a, b \in \mathbf{Z}$). Para cada $x = a + ib \in A$, se define $N(x) = a^2 + b^2$.

a) Probar que $N(xy) = N(x)N(y)$ ($x, y \in A$). Deducir los elementos invertibles de A .

b) Si $x \in A$ y si $N(x)$ es un entero primo, probar que x es un elemento irreducible de A . ¿Es cierto el recíproco?

c) Sean $x \in A$ e $y \in A^*$. Se escribe $x/y = u + iv$, $u, v \in \mathbf{Q}$. Sean u_0, v_0 enteros tales que $|u - u_0| \leq 1/2$ y $|v - v_0| \leq 1/2$. Probar que se verifica

$$(1) \quad x = y(u_0 + iv_0) + r, \quad \text{con} \quad N(r) < N(y)$$

(división euclídea en A). ¿En qué caso las condiciones (1) determinan de forma única $u_0 + iv_0$ y r ?

d) Deducir de c) que todo ideal de A es principal, y que todo elemento de A se escribe de forma única (salvo para factores invertibles) como producto de elementos invertibles de A (inspirarse en el § IV.4).

***15.** a) Sea A el anillo de los enteros de Gauss (ejercicio 14), y \mathbf{Z} es un subanillo de A . Se designa por p un número primo en \mathbf{Z} . Probar que p es irreducible en A si, y sólo si, no existe ningún par (a, b) de enteros > 0 tales que $p = a^2 + b^2$ (utilizar el ejercicio 14).

b) Deducir que p es irreducible en A si, y sólo si, no se verifica $p \equiv 1 \pmod{4}$. (Utilizar los ejercicios 5 y 14.)

***16.** Sea d un elemento de \mathbf{Z} que no sea cuadrado perfecto. Se designa por $\mathbf{Z}[\sqrt{d}]$ al anillo de los números «complejos» de la forma $a + b\sqrt{d}$ ($a, b \in \mathbf{Z}$). Para cada

$$x = a + b\sqrt{d} \in \mathbf{Z}[\sqrt{d}]$$

se escribe $N(x) = a^2 - db^2$. Se dice que $\mathbf{Z}[\sqrt{d}]$ es euclídeo si para todo $x \in \mathbf{Z}[\sqrt{d}]$ y todo y no nulo de $\mathbf{Z}[\sqrt{d}]$, existen dos elementos q, r de $\mathbf{Z}[\sqrt{d}]$ tales que $x = yq + r$ y

$$|N(r)| < |N(y)|.$$

a) Inspirándose en el ejercicio 14, demostrar que $\mathbf{Z}[\sqrt{2}]$ y $\mathbf{Z}[\sqrt{3}]$ son euclídeos.

b) Si $\mathbf{Z}[\sqrt{d}]$ es euclídeo, todo elemento de $\mathbf{Z}[\sqrt{d}]$ se escribe de forma única (salvo para factores invertibles) como producto de elementos irreducibles, y $\mathbf{Z}[\sqrt{d}]$ es principal.

c) Buscar los elementos invertibles de $\mathbf{Z}[\sqrt{2}]$ y $\mathbf{Z}[\sqrt{3}]$ (cf. ejercicio 14 y ejercicio (II.10)).

d) Probar que $\mathbf{Z}[\sqrt{-5}]$ no es euclídeo, apoyándose en la relación

$$(2 + i\sqrt{5})(2 - i\sqrt{5}) = 3 \cdot 3 = 9.$$

Números complejos y Trigonometría

17. Simplificar las expresiones siguientes:

$$S_1 = \sum_{k=0}^n \cos^3 kx, \quad S_2 = \sum_{k=0}^n k^2 \cos(kx),$$

$$S_3 = \sum_{k=0}^{n-1} (x + e^{2ik\pi/n})^n, \quad S_4 = 1 + \sum_{k=1}^{n-1} \frac{\cos kx}{\cos^k x},$$

$$P = \prod_{k=0}^{n-1} (e^{4ik\pi/n} - 2 \cos \theta e^{2ik\pi/n} + 1).$$

(Faddeev-Sominsky.)

(En S_1 se empezará por linealizar; en S_2 se derivará a series conocidas; en S_3 se utilizará la fórmula del binomio; en P se descompondrá cada uno de los factores y se reordenará el producto.)

$$18. \text{ Si establecemos que } A_{n,p} = \sum_{q=0}^{2n-1} \cos^{2p} \left(x + \frac{q\pi}{2n} \right), \text{ probar que, } A_{n,p} = \frac{2n \binom{2}{p}}{2^{2p}}.$$

(Utilizar las fórmulas de Euler y del binomio.) Deducir el valor de la integral $\int_0^\pi \cos^{2p} x \, dx$.
(Utilizar las sumas de Riemann.)

19. Lugar geométrico del punto M de afijo z tal que los puntos de afijos $1, z, \frac{1}{z}, 1-z$ sean cocíclicos (Utilizar el ejercicio II.20.)

(Ecole supérieure d'Electricité.)

20. Sean z_1, z_2, z_3 las raíces, que suponemos simples, de la ecuación $z^3 + pz + q = 0$ y a_1, a_2 las de la ecuación $3z^2 + p = 0$. Z_1, Z_2, Z_3, A_1, A_2 designan las imágenes respectivas de tales números en el plano complejo y se supone que $p \in \mathbf{R}, q \in \mathbf{R}$.

a) Establecer las relaciones

$$(1) \quad (\overrightarrow{Z_3 Z_1}, \overrightarrow{Z_3 A_1}) = (\overrightarrow{Z_3 A_2}, \overrightarrow{Z_3 Z_2}),$$

$$(2) \quad (\overrightarrow{Z_1 Z_2}, \overrightarrow{Z_1 A_1}) = (\overrightarrow{Z_1 A_2}, \overrightarrow{Z_1 Z_3}),$$

$$(3) \quad (\overrightarrow{Z_2 Z_3}, \overrightarrow{Z_2 A_1}) = (\overrightarrow{Z_2 A_2}, \overrightarrow{Z_2 Z_1}).$$

(Para establecer (1), por ejemplo, se comprobará que el número $\frac{a_1 - z_3}{z_1 - z_3} \cdot \frac{a_2 - z_3}{z_2 - z_3}$ es real, utilizando las relaciones entre los coeficientes y las raíces.)

b) Deducir que existe una cónica tangente a los tres lados del triángulo $Z_1 Z_2 Z_3$, y cuyos focos son A_1, A_2 (se utilizará un teorema de Poncelet). ¿Qué pasa si $A_1 = A_2$?

21. Una homografía $z \mapsto \varphi(z) = \frac{az + b}{cz + d}$ es *parabólica* si admite un único punto doble; no parabólica si admite dos puntos dobles distintos (cf. ejercicio II.21).

a) G_φ designa el grupo engendrado por la homografía φ en el grupo circular. Probar que si φ es parabólica, G_φ es isomorfo a \mathbf{Z} .

b) Si φ es no parabólica y si u, v designan los puntos dobles de φ (tomados en un cierto orden) la razón doble $B_\varphi = B(u, v, z, \varphi(z))$ es independiente de z (ejercicio II.21). Probar que la aplicación $H: \varphi \mapsto B_\varphi$ es un homomorfismo inyectivo de G_φ en el grupo multiplicativo \mathbf{C}^* . Examinar el caso en que φ es de orden finito d , y, en particular, el caso en que $d = 2$ (involuciones).

22. Se conservan las notaciones del ejercicio precedente. Una homografía no parabólica φ se llama *hiperbólica* si $H(G_\varphi) \subset \mathbf{R}^*$, *elíptica* si $H(G_\varphi) \subset U$, en donde U designa el grupo multiplicativo de los números complejos de módulo 1.

a) Caracterizar las transformaciones elípticas e hiperbólicas según el valor de B_φ .

b) Toda transformación no parabólica φ se escribe de manera única en la forma $\varphi = \psi \circ \theta$, en donde ψ es hiperbólica y θ es elíptica, y de modo que $\psi \circ \theta = \theta \circ \psi$.

c) Una homografía cualquiera $\varphi(z) = \frac{az + b}{cz + d}$ se puede escribir

$$(z - l) [\varphi(z) - m] = \frac{bc - ad}{c^2}, \quad \text{con: } l = -\frac{d}{c} \quad \text{y} \quad m = \frac{a}{c}.$$

Sean L, M las imágenes respectivas de l, m en el plano complejo. Probar que la transformada de una recta que pase por L es una recta que pasa por M . Expresar B_φ por medio de u, v, l , y después por medio de u, v, m (en donde u, v designan los puntos dobles de φ). Deducir la posición de u, v respecto de l, m en los dos casos:

a) φ es elíptica.

b) φ es hiperbólica.

Probar que en estos dos casos la recta LM es globalmente invariante.

Deducir que si φ es hiperbólica o elíptica, φ es el producto de una inversión y de una simetría respecto de una recta.

23. a) Se dice que cuatro puntos A, B, C, D del plano complejo, distintos entre sí, forman un *cuadrilátero armónico* si sus afijos a, b, c, d verifican $B(a, b, c, d) = -1$ (notaciones del ejercicio II.20). Probar que, en este caso, si M designa el punto medio de AB , la recta CD es simétrica de la recta CM respecto de las bisectrices de AC y BC . ¿Recíproco?

b) Sean φ, ψ dos homografías no parabólicas tales que $\varphi \circ \psi = \psi \circ \varphi$ (ver ejercicio III.21). Si φ y ψ no poseen los mismos puntos dobles, los dos pares de puntos dobles constituyen un cuadrilátero armónico.

CAPÍTULO IV

1. Calcular $\text{mcd}(P, Q)$ cuando

a) $P = X^6 - 7X^4 + 8X^3 - 7X + 7, Q = 3X^5 - 7X^3 + 3X^2 - 7;$

b) $P = X^6 + X^5 - X^4 - 2X^3 - X^2 + X + 1, Q = X^5 + X^3 - X^2 - 1.$

2. ¿Para qué valores de m el polinomio $P_m = (X+1)^m - X^m - 1$ es divisible por $Q = X^2 + X + 1$? La misma pregunta si $P_m = 1 - X^m + X^{2m} - X^{3m} + X^{4m}$ y $Q = 1 - X + X^2 - X^3 + X^4$ (cuerpo de base: \mathbf{C}).

3. El cuerpo de base es \mathbf{C} . ¿Para qué valores de m , P_m es divisible por Q :

a) ¿Cuando $P_m = (X-1)^{m+2} + X^{2m+1}$ y $Q = X^2 - X + 1$?

b) ¿Cuando $P_m = (X+a+b)^{2m+1} - X^{2m+1} - a^{2m+1} - b^{2m+1}$ y $Q = P_1$?

4. Para $1 \leq r \leq k$ se supone $n_r \equiv r-1 \pmod{k}$.

Probar que $P = X^{n_1} + X^{n_2} + \dots + X^{n_k}$ es divisible por $Q = 1 + X + \dots + X^{k-1}$.

(Faddeev-Sominsky.)

5. Descomponer en factores de primer grado los polinomios siguientes:

a) $X^m - \binom{2m}{2} X^{m-1} + \binom{2m}{4} X^{m-2} + \dots + (-1)^m \binom{2m}{m},$

b) $\binom{2m+1}{1} X^m - \binom{2m+1}{3} X^{m-1} + \binom{2m+1}{5} X^{m-2} + \dots + (-1)^m \binom{2m+1}{2m+1},$

c) $(1+X)^m - e^{2i\pi} (1-X)^m.$

(utilizar la expresión de $\frac{\sin n\alpha}{\sin \alpha}$ en función de $\cotg \alpha$).

6. Sea $P \in \mathbf{R}[X]$ tal que $P(x) \geq 0$ para todo $x \in \mathbf{R}$.

Demostrar que existen $S, T \in \mathbf{R}[X]$ tales que $P = R^2 + S^2$.

(Utilizar la descomposición de P en $\mathbf{C}[X]$.)

7. Sea a un real > 0 . Se supone que existe $P \in \mathbf{R}[X]$ tal que $\cos ax = P(\cos x)$. Probar que $a \in \mathbf{N}^*$.

8. Probar que el polinomio $P(X) - X$ divide al polinomio $P[P(X)] - X$.

(Ecole centrale.)

9. a) Sea K un cuerpo conmutativo, infinito o de característica nula, y sea $P \in K[X]$. Se supone que existe $h \in K$ ($h \neq 0$) tal que, para todo $x \in K$, se verifica: $\tilde{P}(x+h) = \tilde{P}(x)$. Probar que P se reduce a una constante. (Con otras palabras, los únicos polinomios periódicos son las constantes.)

b) Sea K el cuerpo finito $\mathbf{Z}/p\mathbf{Z}$ (p primo > 0), y $P \in K[X]$. Se supone que existe un $h \in K$ ($h \neq 0$) tal que: $P(X+h) = P(X)$ (igualdad formal). Demostrar que existe un $S \in K[X]$ tal que, si ponemos $F(X) = X^p - X$, se verifica $P(X) = S(F(X))$. Recíprocamente, probar que todos los polinomios P , de la forma $P(X) = S(F(X))$ verifican la relación $P(X+h) = P(X)$ para todo $h \in K$.

10. Un polinomio normalizado es un polinomio tal que el coeficiente del término de mayor grado es 1. Se designa por I el intervalo $[-1, +1]$ de \mathbf{R} .

a) Para todo entero $n \geq 1$, existe un polinomio normalizado T (único) tal que

$$\frac{1}{2^{n-1}} \cos(n \arccos x) = T(x) \quad \text{para } x \in I.$$

Calcular $\sup_{x \in I} T(x)$. Clasificar las raíces y los extremos de T en I .

b) Sea P un polinomio normalizado de grado n . Estudiar el número de ceros de $P - T$ por medio de los grafos de P y T y deducir que $\sup_{x \in I} |P(x)| \geq \frac{1}{2^{n-1}}$.

11. Se define $B_0(X) = 1$, $B_1 = X - \frac{1}{2}$.

a) Probar que las relaciones

$$\begin{aligned} B'_n(X) &= nB_{n-1}(X) & \text{para } n \in \mathbf{N}, \\ B_n(0) - B_n(1) &= 0 & \text{para } n \text{ entero } \geq 2, \end{aligned}$$

definen una sucesión única de polinomios de $\mathbf{Q}[X]$.

b) Probar que $B_n(1-X) = (-1)^n B_n(X)$.

Determinar el término de mayor grado de $B_n(X)$.

c) Establecer $B_n(X+1) - B_n(X) = nX^{n-1}$ y deducir a partir de ello una expresión de

$$S_n = 1^p + 2^p + \cdots + n^p.$$

(Ecole polytechnique.)

12. Sean $a_0, a_1, \dots, a_n; b_1, \dots, b_n$ números reales. Se pone

$$f(\theta) = a_0 + \sum_{k=1}^n (a_k \cos k\theta + b_k \operatorname{sen} k\theta). \quad (\theta \in \mathbf{R}).$$

a) Probar que $e^{in\theta} f(\theta)$ es un polinomio P en $e^{i\theta}$ del que se precisará el grado. Sea (u_k) la sucesión de coeficientes de P .

b) Hallar una relación entre u_k y $\overline{u_{2n-k}}$. ¿Qué se puede deducir acerca de las raíces de P ?

c) Se supone que $f(\theta_0) = 0$, y que $(\forall \theta) f(\theta) \geq 0$. Probar que $e^{i\theta_0}$ es una raíz de orden par de P .

d) Deducir de lo que antecede que si $f(\theta) \geq 0$ para $\theta \in [0, 2\pi]$, existe una sucesión c_0, c_1, \dots, c_n de números complejos tal que

$$f(\theta) = \left| \sum_{k=0}^n c_k e^{ik\theta} \right|^2.$$

(Ecole polytechnique.)

13. Sea $P \in \mathbf{C}[X]$ tal que $P(0) \neq 0$, que posea todas sus raíces distintas. Probar que existe un $f \in \mathbf{C}[X]$ tal que

$$f^2(X) - X \equiv 0 \pmod{P(X)}.$$

Para ello, se buscará f en la forma

$$f(X) = \sum_i a_i \varphi_i(X), \text{ en donde } \varphi_i(X) = \prod_{j \neq i} (X - \alpha_j),$$

donde las α_j designan las raíces de P .

(E.N.S. Saint-Cloud, 1968.)

14. Sean $P \in \mathbf{Z}[X]$ y $n \in \mathbf{Z}$, y se pone $m = P(n)$. ($\text{gr}(P) \geq 1$).

a) Probar que para todo $k \in \mathbf{Z}$, $P(n + km)$ es divisible por m .

b) Probar que no existe ningún polinomio $P \in \mathbf{Z}[X]$ no constante tal que, para todo $n \in \mathbf{Z}$, $P(n)$ sea primo.

(Ecole polytechnique.)

***15.** Para cada $k \in \mathbf{N}^*$, sea $P_k(X)$ el polinomio $\frac{1}{k!} (X+1)(X+2) \dots (X+k)$.

a) Probar que, para todo $x \in \mathbf{Z}$, $P_k(x) \in \mathbf{Z}$.

b) Probar que todo polinomio $A \in \mathbf{Q}[X]$ de grado n tal que $(\forall x \in \mathbf{Z}) A(x) \in \mathbf{Z}$, es combinación lineal con coeficientes enteros de P_0, P_1, \dots, P_n (en donde $P_0 = 1$)

c) Si $P \in \mathbf{R}[X]$ verifica: $(\forall x \in \mathbf{Q}, P(x) \in \mathbf{Q})$, ¿se tiene: $P \in \mathbf{Q}[X]$?

***16.** p es un número primo > 0 . Se considera un polinomio

$$P \in \mathbf{Z}[X], \quad P = X^n + a_1 X^{n-1} + \dots + a_n,$$

tal que:

a) para $1 \leq k \leq n$, $a_k \equiv 0 \pmod{p}$;

b) $a_n \not\equiv 0 \pmod{p^2}$,

en donde p designa un número primo > 0 .

Considerando el homomorfismo $\bar{\varphi}: \mathbf{Z}[X] \rightarrow \mathbf{Z}/p\mathbf{Z}[X]$ deducido del homomorfismo canónico $\mathbf{Z} \rightarrow \mathbf{Z}/p\mathbf{Z}$, demostrar que P es irreducible en $\mathbf{Z}[X]$ y, por lo tanto, en $\mathbf{Q}[X]$ (cf. Cap. XIV).

(Criterio de Eisenstein.)

c) Deducir la irreducibilidad, para p primo, del polinomio $1 + X + X^2 + \dots + X^{p-1}$ en $\mathbf{Q}[X]$. (Hágase $X = Y + 1$.)

17. K designa un cuerpo conmutativo de característica nula, $(x_i)_{1 \leq i \leq p}$ es una familia finita de elementos de K todos ellos distintos y $(n_i)_{1 \leq i \leq p}$ es una familia de enteros > 0 tales que $\sum_{i=1}^p n_i = n$ (los enteros $n > 1$ y $p > 1$ son fijos).

a) Hallar una base del espacio vectorial E_i de los polinomios P de grado $\leq n-1$ tales que, para todo $j \neq i$, se verifique

$$P^{(k)}(x_j) = 0 \quad \text{para todo } k \text{ tal que } 0 \leq k \leq n_j - 1.$$

Probar que la aplicación $\varphi_i: E_i \rightarrow K^{n_i}$ tal que $\varphi_i(P) = (P(x_i), P'(x_i), \dots, P^{(n_i-1)}(x_i))$ es una biyección.

b) Sea $(\mu_{i,k})_{1 \leq i \leq p, 0 \leq k \leq n_i-1}$ una familia de elementos de K . Probar que existe un polinomio $P \in K[X]$ único, de grado $\leq n-1$, tal que, para $1 \leq i \leq p$ y $0 \leq k \leq n_i-1$, se verifica $P^{(k)}(x_i) = \mu_{i,k}$.

c) Explicitar P en los casos siguientes:

I) $p = n$ (polinomio de interpolación de Lagrange);

II) $n = 2p$ y para $1 \leq i \leq p$, $n_i = 2$; $\mu_{i,0} = \mu_i$, $\mu_{i,1} = \mu_i$.

(Respuesta al caso II):

$$P(X) = \sum_{i=1}^p \left\{ \prod_{\substack{j \neq i \\ 1 \leq j \leq p}} (X - x_j)^2 \left[\mu_i \prod_{j \neq i} \frac{1}{(x_i - x_j)^2} + (X - x_i) \nu_i \prod_{j \neq i} \frac{1}{(x_i - x_j)^2} \left(1 - 2 \mu_i \sum_{j \neq i} \frac{1}{x_i - x_j} \right) \right] \right\}.$$

18. a) Establecer las relaciones

$$(1) \quad \sum_{k=0}^n (-1)^k \binom{n}{k} (X - k)^n = n!$$

(Proceder por recurrencia sobre n , ya directamente, ya derivando el primer miembro.)

b) Utilizando (1) probar que, para $p < n$, se tienen las relaciones:

$$(2) \quad \sum_{k=0}^n (-1)^k \binom{n}{k} (X - k)^p = 0.$$

c) Deducir de todo ello el valor de la suma $\sum_{p=0}^n \binom{n}{p} P(p)$, en donde P es un polinomio de grado $\leq n$.

*19. (Polinomios ciclotómicos.)

Para todo entero $n \geq 1$ se designa por \mathcal{P}_n al conjunto de las raíces primitivas n -ésimas de 1 en \mathbf{C} y por $\Phi_n(X)$ al polinomio

$$\Phi_n(X) = \prod_{\zeta \in \mathcal{P}_n} (X - \zeta).$$

(a) Φ_n se le llama *polinomio ciclotómico* de orden n .)

a) Establecer la fórmula

$$X^n - 1 = \prod_{d|n} \Phi_d(X).$$

Deducir (por recurrencia sobre n) que $\Phi_n(X) \in \mathbf{Z}[X]$. Calcular el término constante de Φ_n . Probar que, para todo divisor estricto δ de n , $\Phi_n(X)$ divide al polinomio $(X^n - 1)/(X^\delta - 1)$. Calcular $\Phi_{p^\alpha}(X)$ (p primo, $\alpha \geq 1$).

b) Calcular el discriminante Δ de $\Phi_{p^\alpha}(X)$ (p primo, $\alpha \geq 1$). Se recuerda (cf. Cap. VI) que Δ está dado por la fórmula

$$\Delta = (-1)^{m(m-1)/2} \prod_{k=1}^m \Phi'_{p^\alpha}(\theta_k),$$

en donde m es el grado de $\Phi_{p^\alpha}(X)$ y en donde $\theta_1, \dots, \theta_m$ son las raíces de $\Phi_{p^\alpha}(X)$.

*20. Todo cuerpo finito es conmutativo. Sea K un cuerpo finito de característica p . Se designa por I el centro de K , es decir, el subcuerpo de los $x \in K$ tales que $(\forall y \in K) xy = yx$. Se sabe (ver el ejercicio III.10) que existen enteros d y f tales que $\text{card}(K) = p^f$ y $\text{card}(I) = p^d = r$.

a) Probar que d divide a f . En lo sucesivo, escribiremos $e = fd$. La hipótesis « K no es conmutativo» equivale a « $f \neq K$ », por lo tanto a « $e > 1$ ».

b) Se hace operar a K^* sobre sí mismo por la izquierda por medio de la ley $(g, x) \mapsto gxg^{-1}$. ¿Cuáles son las órbitas reducidas a un elemento?

c) Sea Ω el conjunto de las órbitas de K^* . Para todo $\omega \in \Omega$, se elige un elemento fijo $x_\omega \in \omega$ y se designa por S_ω al subgrupo de isotropía de x_ω . Probar que aquí la ecuación de clases se escribe

$$(1) \quad \text{card}(K^*) = \sum_{\omega \in \Omega} \frac{\text{card}(K^*)}{[S_\omega]}.$$

d) Para toda órbita $\omega \in \Omega$ probar que el conjunto $K_\omega = S_\omega \cup \{0\}$ es un subcuerpo de K que contiene f . Deducir que existe un divisor δ_ω de e tal que $\text{card}(K_\omega) = r^{\delta_\omega}$. ¿En qué caso $\delta_\omega < e$? Con la ayuda de (1), deducir la fórmula

$$(2) \quad r^e - 1 = r - 1 + \sum_{\substack{\omega \in \Omega \\ \delta_\omega < e}} \frac{r^e - 1}{r^{\delta_\omega} - 1}.$$

e) Se designa por Φ_e al polinomio ciclotómico de orden e (cf. ejercicio precedente). Probar que si $\delta_\omega < e$, el entero $\Phi_e(r)$ divide a $\frac{r^e - 1}{r^{\delta_\omega} - 1}$.

Deducir, utilizando (2), que $\Phi_e(r)$ divide a $r - 1$. Calculando directamente $\Phi_e(r)$, probar que se obtiene una contradicción. Deducir de ello el resultado enunciado.

CAPÍTULO V

1. α, β, γ designan las raíces de $x^3 + px + q = 0$.

Calcular

$$\sum \frac{\alpha^2}{\beta}, \quad \sum \frac{\alpha^2}{\beta^2 + \gamma^2}, \quad \sum \frac{\beta\gamma}{\alpha^2}, \quad \sum \frac{\alpha^2}{\beta + \gamma}.$$

2. Calcular $\sum x_1^2 x_2 x_3$ para las raíces de $x^5 + px + q = 0$.

3. Calcular:

a) $\sum \alpha^5 \beta$ para las raíces de $x^4 + px^3 + qx^2 + rx + s = 0$.

b) $\sum \frac{\alpha\beta}{\gamma^2}$ para las raíces de la misma ecuación.

4. Calcular $\sum (\alpha^3 - \beta^3)^2$ para las raíces de $x^3 + px^2 + qx + 2 = 0$.

*5. Sean $u, v \in \mathbb{C}$ tales que: $uv = p$, $u^5 + v^5 = -2q$, y sea $\theta = e^{2\pi i/5}$.

a) Probar que las raíces de

$$(1) \quad x^5 - 5px^3 + 5p^2x + 2q = 0$$

son:

$$u + v, \quad \theta u + \theta^4 v, \quad \theta^2 u + \theta^3 v, \quad \theta^4 u + \theta v, \quad \theta^3 u + \theta^2 v.$$

b) Se supone que p y q son reales, $p > 0$. Probar que:

Si $p^5 < q^2$ una sola de las raíces de (1) es real.

Si $p^5 > q^2$, 5 las raíces de (1) son reales.

¿Qué pasa si $p^5 = q^2$?

6. Resolver el sistema

$$x^2 + y^2 + z^2 = 2, \quad x^3 + y^3 + z^3 = 2, \quad x^4 + y^4 + z^4 = 2.$$

(Utilizar las fórmulas de Newton)

7. Resolver en \mathbf{C} el sistema

$$x^2 + y^2 + z^2 = 0, \quad x^4 + y^4 + z^4 = 0, \quad x^5 + y^5 + z^5 = 2.$$

(Ecole des Mines.)

8. Descomponer en factores reales el polinomio $X^{2n} + 1$.

9. Calcular $\prod_{k=1}^{n-1} \sin \frac{k\pi}{n}$ y deducir de ello el valor de la integral $\int_0^{\pi/2} \text{Log}(\sin x) dx$.

(Ecole polytechnique.)

10. Sea $P(X)$ un polinomio con coeficientes reales cuyas raíces sean todas ellas reales y distintas, y sea $\alpha \in \mathbf{R}^*$. Probar que $P^2 + \alpha^2$ tiene todas sus raíces distintas.

(Ecole des Mines.)

(Estudiar los ceros de P' .)

11. Estudiar los ceros reales de los polinomios P_n definidos por

$$P_n(X) = \sum_{p=0}^n \frac{X^p}{p!}.$$

(Ecole polytechnique.)

(Utilizar el teorema de Rolle (cf. tomo 2, Cap. IV) y la relación

$$P'_n(X) = P_{n-1}(X).)$$

*12. Sea $P = x^4 + 4ax^3 + 6bx^2 + 4cx + d$. Si $P^{(3)}$ divide a P' , la ecuación $P = 0$ se puede transformar en una ecuación bicuadrada.

*13. Parametrizar la curva de \mathbf{C}^3 de ecuaciones cartesianas

$$x^2 + y^2 + z^2 = 1, \quad x^3 + y^3 + z^3 = 1.$$

(Se tomará $t = x + y + z$ como parámetro, y se definirá x, y, z como raíces de una ecuación dependiente del parámetro t .)

14. Determinar a, b, c de modo que a, b, c sean soluciones de la ecuación

$$x^3 + ax^2 + bx + c = 0.$$

(Ecole des Mines.)

15. Sea $P(z) = z^4 + az^3 + bz^2 + cz + d$ un polinomio de grado 4 con coeficientes en \mathbf{C} . Demostrar la equivalencia de las propiedades a), b), c) dadas a continuación.

a) Las raíces de P forman un paralelogramo.

b) Existe un $h \in \mathbf{C}$ tal que, si hacemos $\Phi(U) = P(h + U)$, $\Phi(U)$ es un polinomio bicuadrado, e.d. de la forma

$$\Phi(U) = AU^4 + BU^2 + C.$$

c) P' y P''' tienen una raíz en común.

16. Determinar la constante A de manera que la ecuación $(1+z)^n = A(1-z)^n$ admita por raíces los números

$$i \cotg \left(x + \frac{k\pi}{n} \right), \quad (0 \leq k \leq n-1),$$

en donde x es fijo. Deducir el valor de $P = \prod_{k=0}^{n-1} \cotg \left(x + \frac{k\pi}{n} \right)$.

17. a) Calcular

$$P(x) = \prod_{k=0}^{n-1} \operatorname{sen} \left(x + \frac{k\pi}{n} \right) \quad (x \notin \pi\mathbf{Z})$$

(generalización del ejercicio V.9; se utilizarán las fórmulas de Euler y se partirá de la relación

$$\prod_{k=0}^{n-1} (X - e^{2ik\pi/n}) = X^n - 1).$$

b) Calcular asimismo

$$Q(x) = \prod_{k=0}^{n-1} \cos \left(x + \frac{k\pi}{n} \right) \quad x \notin \frac{\pi}{2} + \pi\mathbf{Z}.$$

***18.** Sean a_1, \dots, a_n números reales tales que

$$(\forall j) \cos a_j \neq 0 \quad \text{y} \quad \cos(a_1 + \dots + a_n) \neq 0.$$

Para $1 \leq j \leq n$ se establece $u_j = \operatorname{tg} a_j$. Probar que se verifica

$$\operatorname{tg}(a_1 + \dots + a_n) = \left(\sum_{2k+1 \leq n} (-1)^k \sigma_{2k+1} \right) : \left(\sum_{2k \leq n} (-1)^k \sigma_{2k} \right)$$

en donde, para todo $p > 0$, σ_p designa la p -ésima función simétrica elemental de las u_j , y en donde $\sigma_0 = 1$.

19. Se consideran dos polinomios con coeficientes complejos:

$$P(X) = b_0 X^n + \dots + b_n = 0 \quad (b_0 \neq 0),$$

de raíces β_1, \dots, β_n , y $Q(X) = c_0 X^m + \dots + c_m = 0$ ($c_0 \neq 0$), de raíces $\gamma_1, \dots, \gamma_m$. Si para todo entero $k > 0$ se tiene

$$\sum_{i=1}^n (\beta_i)^k = \sum_{j=1}^m (\gamma_j)^k.$$

Probar que toda raíz no nula de P es también una raíz de Q del mismo orden. (Utilizar las fórmulas de Newton.)

(E. N. S. Saint-Cloud, Extracto.)

20. Sea A una \mathbf{R} -álgebra unífera sin divisores de cero y sea I su elemento unidad.

a) Probar que un elemento X de A que no es de la forma $X = \lambda I$, con λ real, y que verifica una relación de la forma

$$(1) \quad a_0 I + a_1 X + \cdots + a_n X^n = 0 \quad (a_0, \dots, a_n \in \mathbf{R})$$

verifica también una relación de la forma

$$(2) \quad X^2 + \alpha X + \beta I = 0,$$

en donde α, β son números reales que cumplen $\alpha^2 - 4\beta < 0$.

b) Probar que si A es conmutativo, existen a lo sumo n elementos de A que verifican la relación (1), en donde a_0, \dots, a_n son reales dados.

c) Probar mediante contraejemplos que estos resultados no son ciertos si A es un álgebra unífera cualquiera.

(Se puede tomar para A el álgebra $\mathcal{M}_2(\mathbf{R})$ (ver Cap. IX) y buscar las matrices $X \in \mathcal{M}_2(\mathbf{R})$ para las cuales $X^2 = I$.)

CAPÍTULO VI

1. Calcular $\lambda \in \mathbf{C}$ para que dos de las raíces de $x^4 - 2x^2 + \lambda x + 3 = 0$ tengan por producto 1. Resolver entonces la ecuación.

2. Condiciones entre p, q, r, s para que las imágenes de las raíces de

$$x^4 + 4px^3 + 6qx^2 + 4rx + s = 0$$

formen un cuadrado. Resolver entonces la ecuación.

3. Calcular λ y $\mu \in \mathbf{C}$ para que la ecuación $X^5 + \lambda X^3 + \mu X + 1 = 0$ tenga una raíz triple.

4. Hallar todas las ecuaciones de grado 3 invariantes por la transformación $y = x - 1/x$.

*5. Hallar la ecuación de grado 3 cuyas raíces son $\beta^2 + \gamma^2 - \alpha^2, \gamma^2 + \alpha^2 - \beta^2, \alpha^2 + \beta^2 - \gamma^2$, en donde α, β, γ designan las raíces de $x^3 + px^2 + qx + r = 0$.

6. Hallar la ecuación cuyas raíces son las 6 razones que se pueden formar con las raíces de $x^3 + px + q = 0$.

7. Eliminar x entre $x^3 - 1 = 0$ y $x^3 + px + q = 0$.

8. Eliminar x entre $x^3 - \lambda x^2 - q = 0$ y $x^3 - \lambda x - 3 = 0$.

(Ecole polytechnique.)

9. Sean P y Q polinomios de grados respectivos m y n . ¿Qué relación existe entre la resultante R de $P(X)$ y $Q(X)$, y la resultante S de los polinomios $M(y)$ y $N(y)$ definidos por

$$M(y) = (\gamma y + \delta)^m P\left(\frac{\alpha y + \beta}{\gamma y + \delta}\right), \quad N(y) = (\gamma y + \delta)^n Q\left(\frac{\alpha y + \beta}{\gamma y + \delta}\right);$$

en donde $\alpha, \beta, \gamma, \delta$ son tales que $\alpha\delta - \beta\gamma \neq 0$?

(Se estudiará el cociente S/R).

10. Determinar un polinomio $P(X, Y, Z, T)$ tal que las relaciones $X = A^2, Y = B^2, Z = C^2, T = D^2$ y $A + B + C + D = 0$ impliquen $P(X, Y, Z, T) = 0$.

(Considerar A, B, C, D como raíces de una ecuación de grado 4 que se transformará por medio de $y = x^2$.)

CAPÍTULO VII

1. El cuerpo de base es \mathbf{C} . Descomponer en elementos de primera especie

$$\frac{(2n)!}{x \prod_{k=1}^n (x^2 - k^2)}, \quad \frac{n!}{x(x-1)\dots(x-n)}, \quad \frac{1}{\cos(n \arccos x)}, \quad \frac{1}{x^m(1-x)^n},$$

$$\frac{1}{(x^2 - a^2)^n}, \quad \frac{1}{(x^n - 1)^2}, \quad \frac{1}{(x^2 + a^2)^n}, \quad \frac{x^4 + 1}{(x^2 - x + 1)^n}.$$

2. Descomponer en elementos simples las fracciones siguientes en $\mathbf{R}(X)$:

$$\frac{x^m}{x^{2n+1} - 1}, \quad \frac{x^m}{x^{2n+1} + 1}, \quad \frac{x^{2m}}{x^{2n} + 1} \quad (m < n), \quad \frac{1}{(x^{2n} - 1)^2},$$

3. Sea $F = \varphi/f$ una fracción racional escrita en forma irreducible; si a es un polo doble de F , se escribe: $\frac{\varphi}{f} = \frac{A}{(x-a)^2} + \frac{B}{x-a} + \dots$.

Probar:

$$A = \frac{2\varphi(a)}{f''(a)}, \quad B = \frac{2}{3} \left[\frac{3\varphi'(a)f''(a) - \varphi(a)f^{(3)}(a)}{[f''(a)]^2} \right].$$

Aplicación: Descomponer $\frac{1}{[H(x)]^2}$, en donde x_1, \dots, x_n son todos distintos y

$$H(x) = \prod_{i=1}^n (x - x_i).$$

4. Sea $\alpha \in \mathbf{R}$. Calcular la derivada n -ésima de $\frac{1}{1 - 2x \cos \alpha + x^2}$.

(Ecole polytechnique.)

5. Simplificar

$$(1) \quad \sum_{k=1}^n \frac{X^3}{(X - \alpha_k)^2} = F \quad \text{cuando} \quad \alpha_k = e^{2ik\pi/n}.$$

(A priori se puede escribir $\sum 1/(X - \alpha_k)^2 = P/Q$, y determinar P y Q , o bien desarrollar cada uno de los términos de la suma (1) en serie formal. Asimismo se podrá efectuar el cálculo directo de $(X^n - 1)^2 F$.)

(Ecole polytechnique.)

6. Reducir a mínimo común denominador y simplificar la suma de las fracciones racionales:

$$\sum_{k=0}^{n-1} \frac{X^2 - \alpha_{k-2} X + \alpha_{k+2}}{(X - \alpha_k)^2} \quad (\alpha_k = e^{2ik\pi/n}).$$

(Se puede descomponer en tres sumas de n términos y desarrollar cada término en serie formal, o efectuar un cálculo directo, como en el ejercicio precedente.)

(Ecole polytechnique.)

7. Se designa por

$$f(X) = X^n + p_1 X^{n-1} + \dots + p_n = (X - X_1)(X - X_2) \dots (X - X_n)$$

un polinomio con coeficientes en \mathbf{C} , y se le asocia el polinomio

$$g(Y) = 1 + p_1 Y + p_2 Y^2 + \dots + p_n Y^n = p_n(Y - Y_1)(Y - Y_2) \dots (Y - Y_n)$$

en donde se ha puesto

$$Y_k = \frac{1}{X_k}, \quad (k = 1, 2, \dots, n).$$

a) Establecer la relación (cuyo segundo miembro es una serie formal):

$$\frac{g'(Y)}{g(Y)} = - \sum_{p \geq 0} S_{p+1} Y^p, \quad \text{con} \quad S_m = \sum_{k=1}^n (X_k)^m.$$

b) Deducir un método de cálculo de los S_m en función de los p_i .

c) Deducir también una nueva demostración de las fórmulas de Newton.

8. Sean $P, Q \in \mathbf{R}[X]$ tales que $\text{gr}(P) = p$, $\text{gr}(Q) = q$, $p < q$.

$P(X) = X^p + a_1 X^{p-1} + \dots + a_p$, $Q(X) = X^q + b_1 X^{q-1} + \dots + b_q$, P y Q primos entre sí. Se pone $R = P/Q \in \mathbf{R}(X)$.

a) Hallar el grado del numerador de $R' = \frac{dR}{dX}$, cuando se elige como representante de R' la fracción $\frac{QP' - PQ'}{Q^2}$.

¿En qué caso este representante de R' es irreducible?

b) Probar que existe una sucesión de polinomios P_n tales que la fracción

$$R^{(n)} = \frac{d^n R}{dX^n} \quad \text{sea igual a la fracción} \quad \frac{P_n(X)}{(Q(X))^{n+1}}.$$

Hallar una relación recurrente entre los P_n .

c) Calcular $\text{gr}(P_n)$.

d) Se supone que $q = 2$, y que las raíces de Q no son reales. Probar por recurrencia que P_n admite $(n + p)$ raíces reales.

(Utilizar el teorema de Rolle.)

(Según un ejercicio propuesto en la Ecole polytechnique.)

9. Se identifica el cuerpo \mathbf{C} de los números complejos con el plano \mathbf{R}^2 , asociando a cada $(x, y) \in \mathbf{R}^2$ el complejo $z = x + iy$.

Sea P un polinomio con coeficientes complejos.

a) Si todas las raíces de P se hallan en el mismo semiplano abierto limitado por una recta Δ del plano, las raíces de $P' = \frac{dP}{dX}$ se hallan en este mismo semiplano.

b) Si todas las raíces de P se hallan en un mismo semiplano cerrado limitado por una recta Δ del plano, y si P' tiene una raíz sobre Δ , P tiene asimismo una raíz sobre Δ .

(Nota: para a) y b) se aconseja utilizar la descomposición de P'/P .)

c) Si las raíces de P se hallan en un polígono convexo cerrado del plano, las raíces de P' se hallan en este mismo polígono.

d) Si las raíces de P son simples, y forman el conjunto de vértices de un polígono convexo, las de P' se hallan en el *interior* de dicho polígono.

10. Se hace $u = \operatorname{tg} x$; probar que, si n es un entero de la forma $n = 2m + 1$, existe una fracción $F \in \mathbf{R}(X)$ tal que $\cotg nx = F(n)$ para $x \in \mathbf{R}$ y $x \notin \mathbf{Z}\pi$.

a) Determinar los polos de F (calcúlese el producto

$$P = \prod_{k=0}^{n-1} \cotg \left(x + \frac{k\pi}{n} \right)$$

y véase ejercicio V.8).

b) Descomponer la fracción F en elementos simples en $\mathbf{R}(X)$. Deducir la expresión de $\cotg x$ en función de $\operatorname{tg}(x/n)$ y de

$$\operatorname{sen} \frac{k\pi}{n}, \quad \cos \frac{k\pi}{n}, \quad \operatorname{tg} \frac{k\pi}{n} \quad (1 \leq k \leq n).$$

$$\left(\text{Respuesta: } \cotg x = \frac{(-1)^m}{n \operatorname{tg} \frac{x}{n}} + \sum_{k=1}^{n-1} \frac{(-1)^m n \operatorname{tg} \frac{x}{n}}{\cos^2 \frac{k\pi}{n} \left(n^2 \operatorname{tg}^2 \frac{x}{n} \right) - n^2 \operatorname{sen}^2 \frac{k\pi}{n}} \right).$$

*c) Deducir que para x real no múltiplo de π , se tiene

$$\cotg x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{2x}{x^2 - k^2 \pi^2}.$$

(Hacer $m = 2p$ en (1) y que p tienda a infinito.)

11. Sean a_1, \dots, a_n elementos distintos de un cuerpo conmutativo K , y sea P un polinomio de grado $\leq n-1$, con coeficientes en K .

a) Descomponer en elementos simples la fracción racional $F(X) = \frac{P(X)}{(X-a_1) \dots (X-a_n)}$

b) Dados n elementos cualesquiera b_1, \dots, b_n de K , probar que existe un polinomio y uno solo $P \in K[X]$, de grado $\leq n-1$, que verifica $P(a_i) = b_i$ para $i = 1, 2, \dots, n$. (Polinomio de interpolación de Lagrange.)

CAPÍTULO VIII

1. Sea E_n el espacio vectorial de los polinomios de grado $\leq n$ sobre un cuerpo K , y sean P_0, P_1, \dots, P_n polinomios tales que $\operatorname{grad}(P_k) = k$ ($0 \leq k \leq n$).

Probar que (P_0, P_1, \dots, P_n) es una base de E_n . Ejemplos de tales bases.

2. En el espacio vectorial $E = \mathcal{F}(\mathbf{R}, \mathbf{R})$ se consideran las funciones $f_y : x \mapsto e^{xy}$.

Probar que $(f_y)_{y \in \mathbf{R}}$ es una familia libre en E .

(Utilizar las derivadas o estudiar el comportamiento en el infinito de las combinaciones lineales de elementos de E .)

3. Sea E el espacio vectorial de las aplicaciones de $[0, 1]$ en \mathbf{R} . Para todo $x \in [0, 1]$ y todo $y \in]0, 1[$, se escribe $f_y(x) = \frac{1}{1-xy}$.

Probar que la familia $(f_y)_{y \in]0, 1[}$ es libre en E .

4. E designa el \mathbf{R} -espacio vectorial $\mathcal{F}(\mathbf{R}, \mathbf{R})$. Probar que, en E , son libres las familias siguientes:

a) $(x^n \operatorname{ch} x, x^p \operatorname{sh} x)_{n \geq 0, p \geq 0}$.

b) $(\cos nx)_{n \geq 0}$.

c) $(\operatorname{sen} nx)_{n \geq 0}$.

5. a) Sean $\alpha_1, \alpha_2, \dots, \alpha_n$ reales distintos, y a_0, a_1, \dots, a_n reales no todos nulos.

Por recurrencia sobre n (y utilizando, por ejemplo, el teorema de Rolle) probar que la ecuación $a_0 + a_1 x^{\alpha_1} + \dots + a_n x^{\alpha_n} = 0$ posee sólo un número finito de soluciones > 0 .

b) T designa un conjunto, y E el \mathbf{R} -espacio vectorial $\mathcal{F}(T, \mathbf{R})$. Sea $f: T \rightarrow \mathbf{R}_+^*$ una aplicación tal que $f(T)$ sea una parte infinita de \mathbf{R}_+^* . Probar que la familia $(f^a)_{a \in \mathbf{R}}$ es libre en E .

6. a) En el espacio vectorial $E = \mathcal{F}(\mathbf{R}, \mathbf{R})$ se consideran los elementos f_a definidos por $f_a(x) = |x - a|$ para $x \in \mathbf{R}$. Probar que la familia $(f_a)_{a \in \mathbf{R}}$ es libre en E .

*b) Probar que la familia $f_{a,\alpha} (a \in \mathbf{R} \setminus 2\mathbf{N}, \alpha \in \mathbf{R})$ definida por $f_{a,\alpha}(x) = |x - a|^\alpha$ es libre en E . (Se puede estudiar el grafo de las funciones que son combinaciones lineales de elementos de la familia considerada.)

7. Sean f_1, f_2, \dots, f_n , n aplicaciones de \mathbf{R} en \mathbf{R} , linealmente independientes en el espacio vectorial $\mathcal{F}(\mathbf{R}, \mathbf{R})$ de las aplicaciones arbitrarias de \mathbf{R} en \mathbf{R} . Probar que existen n números reales x_1, x_2, \dots, x_n tales que $\det [f_i(x_j)]_{1 \leq i \leq n, 1 \leq j \leq n} \neq 0$.

(Ecole polytechnique.)

(Razonar por recurrencia sobre n .)

8. Para $x > -1$ se escribe $g(x) = \operatorname{sen} [\operatorname{Log} (1 + x)]$. Las funciones g, g^2, \dots, g^n , ¿son linealmente independientes?

(Ecole polytechnique.)

(Generalizar y comparar con el ejercicio 5.)

9. Sean F, N dos subespacios vectoriales de un mismo espacio vectorial E . ¿Cuál es la condición necesaria y suficiente para que exista un endomorfismo f de E tal que $f(E) = F, f^{-1}(0) = N$?

(Ecole polytechnique.)

10. Sea E un espacio de dimensión finita n sobre \mathbf{C} , y u, v dos endomorfismos de E tales que $vu = 0$.

Hallar una desigualdad que relacione el rango de u y el rango de v . ¿Puede darse la igualdad? (Ejemplos: las proyecciones.)

(Ecole polytechnique.)

11. Sea E un espacio vectorial, L, M, N tres subespacios. ¿Se cumplen las igualdades

$$L \cap (M + N) = (L \cap M) + (L \cap N),$$

$$L \cap (M + (L \cap N)) = (L \cap M) + (L \cap N) ?$$

(Ecole des Mines.)

12. Sea E un \mathbf{R} -espacio vectorial. Se dice que E es precomplejo si se le puede dotar de una estructura de \mathbf{C} -espacio vectorial, tal que E se obtenga a partir de esta estructura, por restricción de los escalares a \mathbf{R} .

a) Probar que E es precomplejo si, y sólo si, existe un endomorfismo φ de E tal que $\varphi^2 + e = 0$ (e designa el endomorfismo identidad).

b) Cuando E es de dimensión finita n , E es precomplejo si, y sólo si, n es par.

(El problema consiste esencialmente en definir de forma coherente la multiplicación por i .)

13. Sea E un espacio vectorial de dimensión finita n , y sea $u \in \mathcal{L}(E)$.

a) Probar que el núcleo y la imagen de u coinciden si, y sólo si, se verifican las siguientes condiciones:

$$(1) \quad u^2 = 0, \quad n = 2 \dim [u(E)].$$

b) Probar que las condiciones (1) equivalen a la existencia de una base en que la matriz de u es de la forma

$$\begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix}$$

en donde A es una matriz cuadrada invertible de orden $p = n/2$.

14. Si f, g son dos endomorfismos de un espacio vectorial de dimensión n , se tiene

$$\text{rang}(f \circ g) \geq \text{rang}(f) + \text{rang}(g) - n.$$

15. E es el espacio vectorial de las funciones indefinidamente derivables, de período 2π , sobre \mathbf{R} , d designa el endomorfismo: $f \mapsto f'$. Determinar $\text{Ker } d$ e $\text{Im } d$. ¿Se verifica $E = \text{Ker } d \oplus \text{Im } d$?
(Ecole polytechnique.)

16. Sea K un cuerpo conmutativo y n un entero ≥ 2 . Para toda permutación $\sigma, \sigma \in \mathfrak{S}_n$, se define el automorfismo T_σ de $E = K_n$ por $T_\sigma(x_1, x_2, \dots, x_n) = (y_1, \dots, y_n)$, con $y_k = x_{\sigma(k)}$ para $1 \leq k \leq n$. Se dice que un subespacio H de E es *invariante* por \mathfrak{S}_n si es globalmente invariante para todos los automorfismos $T_\sigma (\sigma \in \mathfrak{S}_n)$.

a) Probar que los únicos subespacios invariantes de E , distintos de $\{0\}$ y E , son: el hiperplano $x_1 + x_2 + \dots + x_n = 0$, y la recta de ecuaciones $x_1 = x_2 = \dots = x_n$.

b) ¿Cuáles son los subespacios invariantes por el grupo alternado \mathcal{A}_n ?

17. Sea E un espacio vectorial de dimensión finita n , y sea $u \in \mathcal{L}(E)$. Se establece

$$J_n = u^n(E) \quad (n \geq 0), \quad K_n = \text{Ker}(u^n) \quad (\text{núcleo de } u^n) \quad (n \geq 0).$$

Por convenio, $u^0 = e$, aplicación idéntica de E . Las sucesiones J_n y K_n son estacionarias; además, si s es el mínimo de los enteros p tales que $J_p = J_{p+1}$, se tiene: $K_s = K_{s+1}$, y, si $s \geq r$, $K_{r-1} \neq K_r$. Finalmente, E es suma directa de J_s y K_s .

18. Sea E el espacio vectorial de los polinomios con coeficientes reales, de grado $\leq n$.

a) Probar que los polinomios $((X+h)^n)_{h \in \mathbf{R}}$ forman un sistema de generadores de E .

b) Se considera el conjunto \mathcal{E} de los endomorfismos φ de E tales que $\forall h \in \mathbf{R}, \forall P \in E, \varphi[P(x+h)] = [\varphi(P)](X+h)$; \mathcal{E} es una subálgebra de $\mathcal{L}(E)$. Probar en primer lugar que $\dim(\mathcal{E}) \leq n+1$, después, que $\dim(\mathcal{E}) = n+1$.

(Ecole polytechnique.)

(Demostrar que las aplicaciones $P \mapsto P^{(i)}$ derivada de orden i , para $i = 0, 1, \dots, n$, forman una base de \mathcal{E} .)

19. Sea E el espacio vectorial de los polinomios de grado $\leq n-1$ sobre un cuerpo de característica 0. A partir de un polinomio $P \in E$ se forman los $n+1$ polinomios

$$P_0(X) = P(X); \quad P_1(X) = P(X+1), \dots, P_n(X) = P(X+n).$$

a) Determinar una relación de dependencia entre los polinomios precedentes, esto es:

$$\sum_{k=0}^n a_k P_k = 0.$$

(Para ello considerar el endomorfismo $A: P \mapsto P_1$ de E ; calcular sus valores propios y sus vectores propios; observar que $(A-I)^n = 0$, y $(A-I)^{n-1} \neq 0$; desarrollar $(A-I)^n$ por medio de la fórmula del binomio y aplicarlo a P .)

b) Determinar una base (e_1, \dots, e_n) de E para la que se tenga $B(e_i) = e_{i+1}$, con $B = A - I$.

(Respuesta posible:

$$e_n(X) = 1, \quad e_{n-1}(X) = X, \quad e_{n-2}(X) = \frac{X(X-1)}{2} \text{ etc.})$$

(Ecole polytechnique.)

20. Sea E_n el espacio vectorial de los polinomios de grado $\leq n-1$ con coeficientes reales. Se considera la aplicación lineal $U: E_n \rightarrow E_{n+1}$ tal que $U(P) = Q$, en donde Q está definido por $Q(x) = e^{x^2} \frac{d}{dx} (e^{-x^2} P(x))$, $x \in \mathbf{R}$. Hallar el núcleo de U , y la dimensión de la imagen de U , y la matriz de U en la base canónica $1, X, X^2, \dots, X^{n-1}$.

(E.N.S. des Ponts et Chaussées.)

21. Se considera el conjunto de los polinomios P_n homogéneos de grado n con dos variables X, Y .

- Probar que $B_n = \{X^n, X^{n-1}Y, \dots, XY^{n-1}, Y^n\}$ es una base de P_n .
- Sea Δ la aplicación de P_n en P_{n-2} :

$$P \mapsto \Delta(P) = \frac{\partial^2 P}{\partial X^2} + \frac{\partial^2 P}{\partial Y^2}.$$

Propiedades de Δ . Matriz de Δ en la base (B_n) . Calcular $\dim(\text{Ker } \Delta)$.

(Ecole des Mines.)

22. Sea E el \mathbf{C} -espacio vectorial de los polinomios $P \in \mathbf{C}[X, Y]$ tales que $\text{gr}_X(P) \leq 2$, $\text{gr}_Y(P) \leq 2$.

a) ¿Cuál es la dimensión de E ?

b) Probar que la aplicación $u: E \rightarrow E$ definida por $u(P) = \frac{\partial^2 P}{\partial X \partial Y}$ es un endomorfismo. Determinar la dimensión de la imagen y la dimensión del núcleo de u .

(Ecole polytechnique.)

23. E designa el espacio vectorial de los polinomios de grado $\leq n$, con coeficientes en \mathbf{C} . Sea φ el endomorfismo de E tal que $\varphi(P) = P - P'$.

a) Probar que φ es invertible, y expresar φ^{-1} con la ayuda de los endomorfismos $(D^m)_{m \geq 0}$ así definidos: D^0 es la identidad, $D = D^1$ es el operador de derivación, tal que $D(P) = P'$ y $D^m = D \circ D^{m-1}$ para $m \geq 1$.

b) Sean $\lambda_1, \dots, \lambda_p$ números complejos no nulos distintos. Se hace $\varphi_i(P) = P - \lambda_i P'$ para $P \in E$. Estudiar el operador $\varphi_1 \circ \varphi_2 \circ \dots \circ \varphi_p$.

(Según un ejercicio propuesto por la Ecole polytechnique.)

*24. \mathcal{B} designa el \mathbf{R} -espacio vectorial de las sucesiones acotadas de números reales; \mathcal{C} es el subespacio de \mathcal{B} formado por las sucesiones convergentes; \mathcal{P} es el subespacio de las sucesiones periódicas; \mathcal{C}_0 es el subespacio de las sucesiones que tienen límite 0; l^1 es el subespacio de las sucesiones (x_n) tales que $\sum_{n=1}^{\infty} |x_n|$ es convergente; \mathcal{S} es el subespacio de las sucesiones (x_n) tales que $\sum_{n=1}^{\infty} x_n$ converge.

Probar que los espacios cocientes siguientes son de dimensión infinita:

$$\mathcal{B}/\mathcal{P}, \quad \mathcal{P}/\mathcal{C}, \quad \mathcal{C}/\mathcal{C}_0, \quad \mathcal{C}_0/\mathcal{S}, \quad \mathcal{S}/l^1$$

25. Sean K un cuerpo conmutativo cualquiera, y E el espacio vectorial constituido por las sucesiones $a = (a_n)_{n \in \mathbf{N}}$ de elementos de K , cuyos términos son todos nulos salvo un número finito. ¿Cuál es el dual de E ? Utilizar este dual para construir ejemplos que muestren que ciertos teoremas demostrados en este capítulo para espacios de dimensión finita, no se extienden a los espacios de dimensión infinita.

*26. Sea K un cuerpo conmutativo de característica nula y E_p el espacio vectorial formado por los polinomios $P \in K[X_1, \dots, X_n]$, homogéneos de grado p , y del polinomio nulo.

A todo punto $\lambda = (\lambda_1, \dots, \lambda_n) \in K^n$ se asocia la forma lineal T_λ sobre E_p , tal que: $T_\lambda(P) = P(\lambda)$.

a) Probar que existen m puntos $\lambda^1, \dots, \lambda^m \in K^n$ tales que $T_{\lambda^1}, \dots, T_{\lambda^m}$ forman una base del dual E_p^* de E_p . ($m = \dim(E_p)$). A una base de este tipo se le llamará una λ -base de E_p^* .

b) Utilizar a) para mostrar que todo polinomio $P \in E_p$ se puede escribir en la forma $P = \sum_{j=1}^m d_j U_j^p$,

en donde $d_j \in K$, y en donde U_j es un polinomio homogéneo de grado 1 para todo j .

c) Se designa por F_p el K -espacio vectorial de los polinomios $P \in K(X_1, \dots, X_n]$ de grado $\leq p$. Para todo $\lambda \in K^n$, se define la forma lineal \hat{T}_λ en F_p por $\hat{T}_\lambda(P) = P(\lambda)$. Probar que, si $q = \dim(F_p)$, existen $\lambda^1, \dots, \lambda^q \in K^n$ tales que $\hat{T}_{\lambda^1}, \dots, \hat{T}_{\lambda^q}$ forman una base del dual F_p^* de F_p . Establecer que la base dual de $\hat{T}_{\lambda^1}, \dots, \hat{T}_{\lambda^q}$ está formada por polinomios de grado p .

d) Examinar el caso particular $n = 1$.

(Según el problema escrito, Saint-Cloud, 1967.)

La cuestión b) ha sido propuesta en la Ecole polytechnique.

27. a) Sea A una \mathbf{R} -álgebra unífera de dimensión finita n , sin divisores de cero; y sea I su elemento unidad. Probar que a cada $X \in A$ se pueden asociar $n + 1$ números reales a_0, a_1, \dots, a_n

tales que $a_0 I + \sum_{k=1}^n a_k X^k = 0$. Deducir que si X no es de la forma λI , con λ real, existen $\alpha, \beta \in \mathbf{R}$ que satisfacen $\alpha^2 - 4\beta < 0$ y $X^2 + \alpha X + \beta I = 0$ (cf. ejercicio V.20).

b) Se supone $n > 1$, lo que implica la existencia de un elemento X_0 de A que no es de la forma $X_0 = \lambda I$, con λ real. Probar que es posible elegir $\lambda, \mu \in \mathbf{R}$ tales que $J = \lambda I + \mu X_0$ satisfaga $J^2 = -I$. Inversamente, si $J \in A$ verifica $J^2 = -I$, probar que es posible elegir $\lambda, \mu \in \mathbf{R}$ tales que $X_1 = \lambda I + \mu J$ y $X_2 = \lambda I - \mu J$ verifiquen ambos la relación $X^2 + aX + bI = 0$, en donde a, b son dos números reales tales que $a^2 - 4b < 0$.

c) Suponiendo que A es conmutativo, demostrar que la aplicación $(\lambda + i\mu) \mapsto \lambda I + \mu J$ ($\lambda, \mu \in \mathbf{R}$) es un isomorfismo de \mathbf{C} en A y deducir de ello que A es un cuerpo. (Utilizar el ejercicio V.20.)

(De esta manera se ha establecido que una \mathbf{R} -álgebra conmutativa de dimensión finita que constituye un cuerpo es isomorfa a \mathbf{R} o a \mathbf{C} .)

CAPÍTULO IX

1. Calcular la inversa de la matriz

$$M = \begin{bmatrix} 1 & a & a^2 & \dots & a^n \\ 0 & 1 & a & \dots & a^{n-1} \\ \vdots & & & & \vdots \\ \vdots & & & & a \\ 0 & \dots & \dots & \dots & 1 \end{bmatrix}$$

2. Calcular la inversa de la matriz

$$M = \begin{bmatrix} 1 & 2 & 3 & \dots & n \\ 0 & 1 & 2 & 3 & \dots & n-1 \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 & 1 \end{bmatrix}$$

3. Se pone $\omega_k = e^{2ik\pi/n}$ ($0 \leq k \leq n-1$).

Calculando ${}^tM \cdot \overline{M}$, hallar la inversa de la matriz $M = [(\omega_{k-1})^{j-1}]_{\substack{1 \leq k \leq n \\ 1 \leq j \leq n}}$
(E. N. S. I., 1969, escrito.)

4. Inversa de la matriz

$$M = \begin{bmatrix} 1 + a_1 & 1 & \dots & 1 \\ 1 & 1 + a_2 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & \dots & 1 + a_n \end{bmatrix}, \quad (a_1 \neq 0, a_2 \neq 0, \dots, a_n \neq 0).$$

*5. Sean $M \in \mathcal{M}_n(\mathbf{C})$, $M = [x_{ij}]$. Se supone que, para cada $i = 1, 2, \dots, n$, se tiene

$$|x_{ii}| > \sum_{j \neq i} |x_{ij}|.$$

Probar que M es invertible (establecer la imposibilidad de que exista una relación lineal entre las columnas).

6. Las matrices con coeficientes *reales*, de la forma

$$M = \begin{bmatrix} x & -y & -z & -t \\ y & x & -t & z \\ z & t & x & -y \\ t & -z & y & x \end{bmatrix}$$

forman un cuerpo L con las leyes inducidas por $\mathcal{M}_4(\mathbf{R})$. Buscar el centro de L . Calcular M^n ($n \in \mathbf{N}^*$). (Poner M en la forma $M = xI + N$, y buscar las potencias de N . Cf. también el problema n.º 5).

7. Calcular M^{-1} cuando $M = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ 1 & \beta & \beta^2 & \beta^3 \\ 1 & \gamma & \gamma^2 & \gamma^3 \\ 1 & \delta & \delta^2 & \delta^3 \end{bmatrix}$ ($\alpha, \beta, \gamma, \delta$ distintos).
(Ecole polytechnique.)

8. Probar que la inversa de la matriz

$$A = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 1 & \binom{1}{1} & \dots & \dots & \dots & \vdots \\ 1 & \binom{p}{1} & \dots & \binom{p}{q} & \dots & \binom{p}{p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \binom{n}{1} & \dots & \binom{n}{q} & \dots & \binom{n}{n} \end{bmatrix}$$

es

$$A^{-1} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ -1 & \binom{1}{1} & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ (-1)^p \dots (-1)^{p-q} \binom{p}{q} & \dots & \binom{p}{p} & \dots & \dots & 0 \\ (-1)^n \dots (-1)^{n-q} \binom{n}{q} & \dots & \dots & \dots & \dots & \binom{n}{n} \end{bmatrix}$$

(Ecole polytechnique.)

(Cf. ejercicio XI.19.)

*9. Una matriz $M = [a_{ij}] \in \mathcal{M}_n(\mathbf{C})$ se llama *mágica* si existe un $\delta \in \mathbf{C}$ tal que:

$$\text{para } 1 \leq i \leq n, \sum_j a_{ij} = \delta,$$

$$\text{para } 1 \leq j \leq n, \sum_i a_{ij} = \delta,$$

$$\text{y } \sum_i a_{ii} = \sum_i a_{i, n+1-i} = \delta.$$

Las matrices mágicas forman un subespacio vectorial de $\mathcal{M}_n(\mathbf{C})$. Calcular del mismo la dimensión y dar una base.

(Ecole polytechnique.)

10. Sea K un cuerpo conmutativo. Por definición, el *centro* de $M_n(K)$ es el subanillo de las matrices $C \in M_n(K)$ tales que, para todo $X \in M_n(K)$, $CX = XC$.

Probar que el centro de $M_n(K)$ está formado por las matrices λI ($\lambda \in K$).

(Considerar ante todo el caso en que X es diagonal.)

(Ecole polytechnique.)

11. Sea K un cuerpo conmutativo. Para toda matriz $A = [a_{ij}]$ perteneciente a $M_n(K)$, se designa por $\text{Tr}(A) = \sum_{i=1}^n a_{ii}$ a la *traza* de A .

a) Probar que, cualesquiera que sean las matrices $A, B, C \in M_n(K)$, se tiene:

$$\text{Tr}(AB) = \text{Tr}(BA); \quad \text{Tr}(ABC) = \text{Tr}(BCA).$$

b) Sean A, B dos matrices tales que, para toda matriz $X \in M_n(K)$, se verifica:

$$\text{Tr}(AX) = \text{Tr}(BX).$$

Probar que $A = B$.

12. Las notaciones son las mismas que las del ejercicio anterior.

a) Probar que a cada aplicación lineal f de $M_n(K)$ en K se le puede asociar una matriz F única tal que $f(A) = \text{Tr}(AF)$ para toda matriz $A \in M_n(K)$.

b) Probar que toda aplicación lineal f de $M_n(K)$ en K que verifica $f(AB) = f(BA)$ cualesquiera que sean $A, B \in M_n(K)$ es de la forma $f(A) = \lambda \text{Tr}(A)$, en donde $\lambda \in K$.

(Utilizar los ejercicios precedentes.)

13. a) Sea $u \in \mathcal{L}(\mathbf{C}^n)$ un endomorfismo que no se reduzca a una homotecia. Probar que existe una base (e_1, \dots, e_n) de \mathbf{C}^n tal que $u(e_1) = e_2$.

b) Sea $A \in \mathcal{M}_n(\mathbf{C})$ tal que $\text{Tr}(A) = 0$. Probar, utilizando a) que A es semejante a una matriz cuya diagonal únicamente contiene ceros (razonar por recurrencia).

14. Para toda matriz nilpotente $N \in M_n(\mathbf{C})$ se define

$$(1) \quad \exp(N) = \sum_{p=0}^{\infty} \frac{1}{p!} N^p$$

(la suma (1) se reduce a un número finito de términos).

c) Sean N_1, N_2 dos matrices nilpotentes tales que $N_1 N_2 = N_2 N_1$; probar que la matriz $N_1 + N_2$ es nilpotente y que

$$\exp(N_1 + N_2) = \exp(N_1) \cdot \exp(N_2).$$

Deducir que $\exp(N)$ es invertible, cualquiera que sea la matriz nilpotente N .

b) Suponiendo en todo momento que N es nilpotente, se define

$$(2) \quad \log(I + N) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} N^n.$$

Probar que $\log(\exp N) = N$ y $\exp[\log(I + N)] = I + N$.

(En el Curso de Análisis (tomo II) un estudio de la convergencia de las series (1) y (2) permitirá extender las aplicaciones $N \mapsto \exp(N)$ y $N \mapsto \log(I + N)$ al caso de matrices no nilpotentes.)

CAPÍTULO X

1. Calcular el determinante

$$\begin{vmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{vmatrix}$$

(Descomponer en un producto de factores lineales.)

2. Calcular el determinante

$$\Delta_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{s-1} & x_1^{s+1} & \dots & x_1^n \\ \vdots & & & & & & & \\ 1 & x_n & x_n^2 & \dots & x_n^{s-1} & x_n^{s+1} & \dots & x_n^n \end{vmatrix}$$

3. Para cada entero $k = 1, 2, \dots, n-1$, se da un polinomio f_k , de grado k , de la forma

$$f_k(X) = X^k + \sum_{p=0}^{k-1} a_{k,p} X^p$$

Establecer la relación

$$\begin{vmatrix} 1 & f_1(x_1) & \dots & f_{n-1}(x_1) \\ \vdots & \vdots & & \vdots \\ 1 & f_1(x_n) & \dots & f_{n-1}(x_n) \end{vmatrix} = \begin{vmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n^{n-1} \end{vmatrix} \quad (\text{Faddeev-Sominsky}).$$

Deducir de ello un cálculo del determinante

$$\Delta_{n+1} = \begin{vmatrix} 1 & \cos a_1 & \dots & \cos(na_1) \\ \vdots & \vdots & & \vdots \\ 1 & \cos a_{n+1} & \dots & \cos(na_{n+1}) \end{vmatrix}.$$

4. Calcular los determinantes siguientes:

$$A_{2n} = \begin{vmatrix} a & 0 & \dots & 0 & b \\ 0 & a & & & 0 \\ & & a & b & \\ & & b & a & \\ 0 & & & & 0 \\ b & 0 & \dots & 0 & a \end{vmatrix}, \quad B_n = \begin{vmatrix} 2 \cos \theta & 1 & 1 & 0 & \dots & 0 \\ & 1 & 2 \cos \theta & 1 & 0 & \dots & 0 \\ & 0 & 1 & \ddots & \ddots & \ddots & \vdots \\ & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ & & & & & & 1 \\ 0 & 0 & & & & 1 & 2 \cos \theta \end{vmatrix}$$

$$C_n = \begin{vmatrix} \cos \theta & 1 & 0 & \dots & 0 \\ 1 & 2 \cos \theta & 1 & \ddots & \vdots \\ 0 & 1 & 2 \cos \theta & 1 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & 1 & 2 \cos \theta \end{vmatrix} \quad (\text{Proceder por recurrencia.})$$

*5. Demostrar las relaciones:

$$\begin{vmatrix} 1+x_1 & 1+x_1^2 & \dots & 1+x_1^n \\ \vdots & \vdots & & \vdots \\ 1+x_n & 1+x_n^2 & \dots & 1+x_n^n \end{vmatrix} = \prod_{i < j} (x_j - x_i) [2 x_1 \dots x_n - (x_1 - 1)(x_2 - 1) \dots (x_n - 1)]$$

$$\begin{vmatrix} 1 & x & x^2 & \dots & x^n \\ 1 & 2x & 3x^2 & \dots & (n+1)x^n \\ 1 & 2^2x & 3^2x^2 & \dots & (n+1)^2x^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2^{n-1}x & 3^{n-1}x^2 & \dots & (n+1)^{n-1}x^n \\ 1 & y & y^2 & \dots & y^n \end{vmatrix} = \left(\prod_{k=1}^{n-1} k! \right) x^{n(n-1)/2} (y-x)^n.$$

6. Calcular el determinante

$$\Delta_n = \begin{vmatrix} 1 & 2 & 3 & \dots & n \\ n & 1 & 2 & 3 & \dots & n-1 \\ n-1 & n & 1 & 2 & \dots & . \\ \vdots & & \ddots & \ddots & \ddots & \ddots \\ 2 & 3 & \dots & \dots & n & 1 \end{vmatrix}.$$

(Se calculará el producto $\prod_{\omega \in U_n} (1 + 2\omega + \dots + n\omega^{n-1})$ en donde U_n designa el grupo de las raíces n -ésimas de la unidad.)

7. Sea $B = [b_{ij}]$ una matriz cuadrada de orden $n+1$ con coeficientes en \mathbf{C} . Para

$$1 \leq i \leq n+1,$$

se hace $B_i(x) = \sum_{k=1}^{n+1} b_{ik} x^{k-1}$, y se designa por $X = [a_{ij}]$ a la matriz tal que $a_{ij} = x_j^{i-1}$ para $i, j = 1, 2, \dots, n+1$.

a) Calcular BX .

b) Aplicación: Por medio de los polinomios $B_i(x) = (x+i)^n$, deducir el valor del determinante

$$\Delta_n = \begin{vmatrix} 1^n & 2^n & \dots & (n+1)^n \\ 2^n & 3^n & \dots & . \\ \vdots & & \ddots & \vdots \\ (n+1)^n & (n+2)^n & \dots & (2n+1)^n \end{vmatrix}.$$

c) Calcular asimismo:

$$D_n = \begin{vmatrix} 0 & (-1)^{n-1} & \dots & (1-n)^{n-1} \\ 1^{n-1} & \ddots & \ddots & 0 & (-1)^{n-1} & \dots & (2-n)^{n-1} \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ (n-1)^{n-1} & \dots & \dots & \dots & \dots & \dots & 0 \end{vmatrix}$$

*8. Probar que:

$$\det \left[\frac{1}{a_i + b_j} \right]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} = \frac{\prod_{i < j} (a_j - a_i)(b_j - b_i)}{\prod_{i,j} (a_i + b_j)}.$$

9. Sea $M(t) \in \mathcal{M}_n(\mathbf{R})$ una matriz cuadrada invertible cuyos términos son funciones derivables del parámetro $t \in \mathbf{R}$.

- a) Escribir la fórmula de Taylor-Young de orden 1 para $M(t)$ en un entorno de t_0 .
 b) I_n designa la matriz unidad de orden n . Se define

$$L(h) = \det [I_n + h[M(t_0)] M'(t_0)] .$$

$L(h)$ es un polinomio en h . Explicítese el término constante y el término de grado 1.

- c) Deducir la fórmula

$$\frac{d}{dt} [\det (M(t))] = \det (M(t)) \cdot \text{Tr} \left[M^{-1}(t) \frac{d}{dt} (M(t)) \right] .$$

- d) Aplicación. Si $A \in \mathcal{M}_n(\mathbf{R})$, calcular $\varphi(t) = \det(\exp(tA))$.

(Se recuerda que, para toda $M \in \mathcal{M}_n(\mathbf{C})$, $\exp(M) = \sum_{n \geq 0} \frac{M^n}{n!}$).

N. B. La derivada de una matriz $A = [a_{ij}(t)]$ es la matriz $A' = \left[\frac{d}{dt} a_{ij}(t) \right]$. No existe ninguna relación simple entre $\det(A')$ y $\frac{d}{dt} [\det(A)]$. En el Curso de Análisis se halla otra fórmula para expresar la derivada de un determinante.

Remitirse al Curso de Análisis para las cuestiones de convergencia.

10. Calcular los determinantes siguientes:

$$D = \begin{vmatrix} 1 & x & x^2 & x^3 \\ x^3 & x^2 & x & 1 \\ 1 & 2x & 3x^2 & 4x^3 \\ 4x^3 & 3x^2 & 2x & 1 \end{vmatrix}, \quad \Delta = \begin{vmatrix} 1 & x & x^3 & x^3 & x^4 \\ 1 & 2x & 3x^2 & 4x^3 & 5x^4 \\ 1 & 4x & 9x^2 & 16x^3 & 25x^4 \\ 1 & y & y^2 & y^3 & y^4 \\ 1 & 2y & 3y^2 & 4y^3 & 5y^4 \end{vmatrix}$$

(Respuesta: $\Delta = 2x^3y(x-y)^6$.)

11. a) Sea K un cuerpo conmutativo. Si el producto PQ de los polinomios $P, Q \in K[X_1, \dots, X_n]$ es homogéneo, cada uno de los polinomios P, Q lo es (ver Cap. IV).

- b) Utilizar a) para demostrar que el polinomio $\Delta_n((X_{ij}))$ con n^2 variables

$$\Delta_n((X_{ij})) = \begin{vmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ \vdots & & & \\ X_{n1} & \dots & \dots & X_{nn} \end{vmatrix} \quad \text{es irreducible en } K[(X_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}] .$$

Se razonará por reducción al absurdo y se utilizará la homogeneidad respecto de las filas o de las columnas.

- c) Probar asimismo que el polinomio con $\frac{n(n+1)}{2}$ variables $(X_{ij})_{1 \leq i \leq j \leq n}$ definido por

$S_n((X_{ij})) = \det([y_{ij}])$ en donde se ha puesto: $y_{ij} = X_{ij}$ para $i \leq j$ e $y_{ij} = X_{ji}$ para $i > j$, es irreducible en $K[(X_{ij})_{1 \leq j \leq i \leq n}]$.

12. a) Probar que las matrices de la forma

$$A = \begin{bmatrix} u & -\bar{v} \\ v & \bar{u} \end{bmatrix} \quad (u, v \in \mathbf{C})$$

constituyen una \mathbf{R} -álgebra K . La aplicación $\varphi: A \mapsto \varphi(A) = (|u|^2 + |v|^2)^{1/2}$ ¿constituye una norma en K (considerado como espacio vectorial)?

b) Probar que K es un cuerpo no conmutativo (cuerpo de los cuaterniones) (cf. problema n.º 5).

CAPÍTULO XI

1. Calcular los valores propios de las matrices siguientes:

$$\begin{bmatrix} 0 & x & \dots & x \\ y & 0 & \dots & . \\ . & & & . \\ . & & & x \\ y & \dots & y & 0 \end{bmatrix} \quad (\text{orden } n); \quad \begin{bmatrix} 0 & -1 & 0 & \dots & 0 \\ -1 & 0 & -1 & \dots & . \\ 0 & & & & . \\ . & & & & -1 \\ 0 & \dots & \dots & -1 & 0 \end{bmatrix} \quad (\text{orden } n);$$

$$\begin{bmatrix} a_1 & a_2 & \dots & a_n \\ a_n & a_1 & \dots & a_{n-1} \\ \vdots & & \vdots & \\ a_2 & a_3 & \dots & a_n & a_1 \end{bmatrix};$$

2. Valores propios y vectores propios de A :

$$A = \begin{bmatrix} x & -y & -z & -t \\ y & x & -t & z \\ z & t & x & -y \\ t & -z & y & x \end{bmatrix}, \quad x, y, z, t \text{ son números reales no todos nulos}$$

3. ¿Condición para que $M = \begin{bmatrix} 0 & \dots & 0 & \alpha_1 \\ . & & \alpha_2 & . \\ 0 & \dots & . & . \\ \alpha_n & 0 & \dots & 0 \end{bmatrix}$ sea diagonalizable?

4. Sea \mathcal{M} el conjunto de las matrices cuadradas de orden n sobre \mathbf{R} con coeficientes reales, de la forma

$$P = [p_{ij}] \quad (i, j = 1, 2, \dots, n)$$

que verifican:

$$(\forall i, j) \quad 0 \leq p_{ij} \leq 1 \quad \text{y} \quad (\forall i) \quad \sum_{j=1}^n p_{ij} = 1.$$

a) Si $P, Q \in \mathcal{M}$, se tiene $PQ \in \mathcal{M}$.

b) Demostrar que toda matriz $P \in \mathcal{M}$ admite el valor propio 1.

(Ecole polytechnique.)

5. Sea E un \mathbf{C} -espacio vectorial de dimensión 6. Buscar los endomorfismos f de E que verifican

$$(f^2 - f + 3I)(f - 2I)^2 = 0, \text{ con } (f - 2I)^2 \neq 0 \text{ y } (f - 2I)(f^2 - f + 3I) \neq 0,$$

en donde I designa la aplicación idéntica de E . ¿Son diagonalizables estos endomorfismos?

(Ecole polytechnique.)

6. Construir un endomorfismo de \mathbf{C}^n que admita un polinomio característico P dado y un polinomio minimal Q dado (Q es un divisor de P).

7. m designa un número real, y se considera la matriz

$$A_m = \begin{pmatrix} m & 1 & 1 \\ 1 & m & 1 \\ 1 & 1 & m \end{pmatrix}.$$

a) Determinar sus valores propios y los vectores propios asociados. Probar que A_m es diagonalizable.

b) Discutir, según los distintos valores de m , el rango de A_m . Determinar, cuando exista, la matriz inversa de A_m .

c) En el caso de ser A_m no invertible, determinar el núcleo y la imagen del endomorfismo asociado.

8. Sea E un espacio vectorial sobre \mathbf{R} , y $f \in \mathcal{L}(E)$ tal que $f^2 = f$. Condición que debe verificar $t \in \mathbf{R}$ para que $I + tf$ sea invertible. Calcular su inversa.

(Ecole polytechnique.)

9. Sea E un espacio vectorial de dimensión finita n sobre \mathbf{C} , y sea $f \in \mathcal{L}(E)$ tal que $f^2 = -I_E$. Hallar los valores propios y los vectores propios de f . Probar que E es diagonalizable.

(Ecole polytechnique.)

10. A es la matriz de un cierto endomorfismo de un espacio vectorial de dimensión n tal que $A^3 = I$. Probar que A es diagonalizable en \mathbf{C} .

(Ecole polytechnique.)

11. Vectores propios y valores propios del endomorfismo A de \mathbf{R}^3 definido por $A(\overrightarrow{W}) = \overrightarrow{V} \wedge \overrightarrow{W}$, en donde \overrightarrow{V} es un vector dado de \mathbf{R}^3 .

(Considerar \mathbf{R}^3 como espacio complejo a fin de hallar vectores y valores propios complejos.)

12. Sea $A \in \mathcal{M}_n(K)$ y sea k el grado del polinomio minimal de A .

a) Probar que el subespacio vectorial de $\mathcal{M}_n(K)$ engendrado por las matrices $A^m (m \in \mathbf{N})$ es de dimensión k .

b) Suponiendo que A es invertible, probar que el subespacio de $\mathcal{M}_n(K)$ engendrado por las matrices $A^m (m \in \mathbf{Z})$ es también de dimensión k .

13. Determinar el polinomio minimal de la matriz

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Deducir un cálculo rápido de A^{-1}, A^3, A^{-3} .

14. Sea E_n el espacio vectorial de los polinomios de grado $\leq n$ sobre \mathbf{R} , y sea φ el endomorfismo de E_n definido por

$$\varphi(P) = (X^2 - 1) P'' + (2X + 1) P'.$$

Hallar la matriz de φ en la base $(1, X, \dots, X^n)$, y los valores propios. Probar que los vectores propios asociados a $\lambda_p = p(p+1)$ son polinomios de grado p .

(Ecole polytechnique.)

15. Sean M, A dos matrices cuadradas de orden n con coeficientes en \mathbf{C} tales que $MA = AM$, y se supone que M tiene todos sus valores propios distintos.

- Probar que todo vector propio de M es un vector propio de A .
- Probar que A es de la forma

$$A = \alpha_0 I_n + \alpha_1 M + \dots + \alpha_{n-1} M^{n-1} \quad (\alpha_0 \dots \alpha_n \in \mathbf{C}).$$

(Utilizar una base de vectores propios.)

(Según un problema propuesto en la Ecole polytechnique.)

16. a) ¿Qué condición debe satisfacer el endomorfismo $f \in \mathcal{L}(E)$ para que exista un vector $x_0 \in E$ tal que los n vectores $x_0, f(x_0), \dots, f^{n-1}(x_0)$ constituyan una base de E ?

- Sea $g \in \mathcal{L}(E)$ definido por

$$g(x) = a_0 x + a_1 f(x) + \dots + a_{n-1} f^{n-1}(x).$$

Probar que $g(x_0) = 0$ implica $g = 0$.

- Probar que si el endomorfismo h conmuta con f , entonces h es combinación lineal de $I, f, f^2, \dots, f^{n-1}$.

(El mismo resultado que en el ejercicio precedente.)

17. Se designa por (e_1, e_2, e_3, e_4) a la base canónica de \mathbf{R}^4 , considerado como espacio vectorial sobre \mathbf{R} , y se considera el endomorfismo u de \mathbf{R}^4 definido por:

$$\begin{aligned} u(e_1) &= e_1 + e_3 + e_4, \\ u(e_2) &= -e_1 - e_3 - e_4, \\ u(e_3) &= 2e_1 + e_2 + e_3 + e_4, \\ u(e_4) &= -2e_1 - e_2. \end{aligned}$$

1) Calcular el polinomio característico $P_u(x)$ de u y probar que $\lambda_1 = 0$ y $\lambda_2 = 1$ son raíces de $P_u(x) = 0$.

2) Probar que existen dos subespacios vectoriales suplementarios E_1 y E_2 de \mathbf{R}^4 estables para u y tales que la restricción de $(u - \lambda_i I)$ a E_i es nilpotente para $i = 1$ e $i = 2$. (I designa el endomorfismo identidad de \mathbf{R}^4 .)

3) Sea u_i la restricción de u a E_i ($i = 1, 2$). Determinar una base de E_i tal que la matriz de u_i respecto de dicha base sea de la forma

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{para } i = 1; \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{para } i = 2.$$

(MP2, París.)

18. Sea A una matriz nilpotente de $M_n(\mathbf{R})$ no nula.

1) Sea I_n la matriz identidad de $M_n(\mathbf{R})$. Probar que $I_n - A$ no es diagonalizable.

2) Si B es una matriz diagonalizable de $M_n(\mathbf{R})$ que posee todos sus valores propios iguales, probar que $A + B$ no es diagonalizable.

3) En $M_2(\mathbf{R})$ hallar una matriz A nilpotente y una matriz B diagonal tales que $A + B$ sea diagonalizable.

(MP2, París.)

19. A cada número real t se le asocia el vector $X(t)$ de \mathbf{R}^n de componentes $(1, t, t^2, \dots, t^{n-1})$.

a) A cada sistema de n vectores Y_0, Y_1, \dots, Y_{n-1} de \mathbf{R}^n se le asocia la aplicación $t \mapsto Y(t)$, de \mathbf{R} en \mathbf{R}^n , definida por

$$Y(t) = \sum_{k=0}^{n-1} t^k Y_k.$$

Probar que existe un único endomorfismo A , independiente de Y_0, Y_1, \dots, Y_{n-1} que verifica: $A \cdot X(t) = Y(t)$ para $t \in \mathbf{R}$.

b) Escribir la matriz asociada al endomorfismo A_u definido por $A_u \cdot X(t) = X(t + u)$, en donde u designa un número real dado, y probar que los endomorfismos A_u constituyen un grupo multiplicativo.

c) Formar la matriz A (correspondiente a $u = 1$) y utilizar b) para tener su inversa

d) ¿Qué condiciones debe verificar el polinomio $P_p(\lambda) = \sum_{k=0}^p a_k \lambda^k$ con coeficientes en \mathbf{C} para

que se verifique $P_p(A_1) = 0$? Probar que se tiene $p \geq n$; que P_n existe y que es único salvo para un coeficiente. Determinar P_n sin cálculo.

(MP2, París.)

20. Sea un polinomio $P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_0$ con coeficientes complejos y sea u_P el endomorfismo de \mathbf{C}^n definido por:

$$(1) \quad \begin{cases} u_P(e_i) = e_{i+1} & 1 \leq i \leq n-1 \\ u_P(e_n) = -(a_0 e_1 + a_1 e_2 + \dots + a_{n-1} e_n) \end{cases}$$

en donde (e_1, e_2, \dots, e_n) es la base canónica de \mathbf{C}^n .

1) Calcular en función de P , polinomio característico de u_P .

2) Comprobar $P(u_P) = 0$.

3) Probar, con la ayuda de las relaciones (1), que todo polinomio S de $\mathbf{C}[X]$ tal que $\text{gr } S < n$ y $S(u_P) = 0$ es nulo.

4) De 2) y 3) deducir que todo polinomio Q de $\mathbf{C}[X]$ tal que $Q(u_P) = 0$ es un múltiplo de P .

5) Si u_P es diagonalizable, probar que P carece de raíces simples.

(MP2, París.)

6) En el caso general, probar que P es el polinomio minimal de u_P .

21. Sean $A \in \mathcal{M}_n(\mathbf{C})$ y $S \in \mathbf{C}[X]$. Si el polinomio característico de A es

$$P(X) = (-1)^n (X - \lambda_1)(X - \lambda_2) \dots (X - \lambda_n),$$

el de $S(A)$ es

$$(-1)^n [X - S(\lambda_1)] \dots [X - S(\lambda_n)].$$

22. Sean $A \in \mathcal{M}_n(\mathbf{C})$, $B \in \mathcal{M}_n(\mathbf{C})$. Probar que AB y BA tienen el mismo polinomio característico. (Empezar considerando el caso en que A es invertible, y en el caso general, substituir A por $A + \lambda I_n$. Cf. también el ejercicio 29.)

23. a) Sea $A \in \mathcal{M}_n(\mathbf{C})$. Probar que A es nilpotente si, y sólo si, se verifica $\text{Tr}(A^k) = 0$ para $1 \leq k \leq n$. Calcular entonces $\det(I_n + A)$.

(Expresar la matriz A en forma triangular y utilizar el § V.2.)

b) Si A es nilpotente, $M \in \mathcal{M}_n(\mathbf{C})$ y $AM = MA$, probar que

$$\det(M + A) = \det(M).$$

(Empezar por el caso en el que M es invertible.)

24. a) Sea $A \in \mathcal{M}_n(\mathbf{R})$ una matriz cuyo polinomio minimal es $X^2 + aX + b = 0$.

¿Qué condiciones deben verificar a y b para que las matrices $\lambda I + \mu A$ ($\lambda, \mu \in \mathbf{R}$) constituyan un cuerpo? Probar que este cuerpo es isomorfo a \mathbf{C} .

b) Inversamente, sea K una subálgebra de $\mathcal{M}_n(\mathbf{R})$ que sea cuerpo conmutativo.

Utilizando los resultados del ejercicio VIII.26 probar que K es el cuerpo de las matrices λI ($\lambda \in \mathbf{R}$) o uno de los cuerpos obtenidos en a).

25. Sea (X_n) ($n \in \mathbf{N}$) una sucesión de vectores de \mathbf{R}^p cuyas componentes $X_{n,i}$ ($i = 1, 2, \dots, p$) verifican las relaciones de recurrencia.

$$X_{n+1,i} = \frac{1}{p-1} \sum_{j \neq i} X_{n,j}.$$

Formar la matriz A tal que $X_{n+1} = AX_n$ cualquiera que sea n . Comprobar que el polinomio minimal de A es

$$(p-1)X^2 - (p-2)X - 1 = 0.$$

Deducir la expresión de las sucesiones $f(n)$ y $g(n)$ tales que $A^n = f(n)A + g(n)I_p$. Determinar $\lim_{n \rightarrow \infty} A_n$ y $\lim_{n \rightarrow \infty} X_n$ en función de X_0 .

26. E designa un espacio vectorial de dimensión n sobre un cuerpo conmutativo y se considera un endomorfismo diagonalizable fijo $f \in \mathcal{L}(E)$. Para toda forma φ bilineal definida en E , se define $\psi = T_\varphi$, en donde

$$\psi(x, y) = \varphi[f(x), y] + \varphi[x, f(y)].$$

Probar que T es un endomorfismo diagonalizable del espacio de las formas bilineales en E . (Utilizar una base de E , conveniente.)

27. Para cada permutación $\sigma \in \mathfrak{S}_n$, sea T_σ el endomorfismo de \mathbf{C}^n definido por

$$T_\sigma(x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Probar que T_σ es diagonalizable y dar una base de vectores propios.

(Empezar considerando el caso en el que σ es una permutación circular y, a continuación, utilizar el teorema (II.8.5).)

28. Sean $A \in \mathcal{M}_p(\mathbf{C})$ y $B \in \mathcal{M}_q(\mathbf{C})$ dos matrices cuadradas tales que, para todo

$$k \in \mathbf{N}, \quad \text{Tr}(A^k) = \text{Tr}(B^k).$$

Demostrar que A y B tienen los mismos valores propios no nulos, con el mismo orden de multiplicidad. Comparar sus polinomios característicos.

(Utilizar el ejercicio V.19)

(E. N. S. Saint-Cloud.)

29. Sean A, B dos matrices rectangulares: $A \in \mathcal{M}_{p,q}(\sigma)$ y $B \in \mathcal{M}_{q,p}(\mathbf{C})$, tales que los productos AB y BA estén definidos.

- Comparar $\text{Tr}(AB)$ y $\text{Tr}(BA)$; y, en general, comparar $\text{Tr}((AB)^k)$ y $\text{Tr}((BA)^k)$ para $k \in \mathbf{N}^*$.
- Demostrar que AB y BA tienen los mismos valores propios no nulos, con los mismos órdenes de multiplicidad (ver ejercicio precedente).
- Si $p = q$, probar que AB y BA tienen el mismo polinomio característico.

(E. N. S. Saint-Cloud.)

30. Sea Φ un polinomio homogéneo de grado n respecto de las n^2 incógnitas (X_{ij}) , $1 \leq i \leq n$, $1 \leq j \leq n$. Identificamos el espacio vectorial $\mathcal{M}_n(\mathbf{C})$ con \mathbf{C}^{n^2} . Suponemos que Φ es *multiplicativo*, e.d. que verifica:

$$\Phi(X \cdot Y) = \Phi(X) \Phi(Y) \quad \text{para } X, Y \in \mathcal{M}_n(\mathbf{C}).$$

- Probar que $\Phi(I_n) = 1$, y que, si X es regular, se tiene:

$$\Phi(X) \neq 0.$$

- Determinar la restricción de Φ a las matrices diagonales.

- Sea $X \in \mathcal{M}_n(\mathbf{C})$. Establecer que, si X posee todos sus valores propios distintos, los polinomios $D(\lambda) = \det(X + \lambda I_n)$ y $F(\lambda) = \Phi(X + \lambda I_n)$ coinciden; deducir en este caso:

$$\Phi(X) = \det(X).$$

- Probar que $\Phi(X) = \det(X)$ para todo $X \in \mathcal{M}_n(\mathbf{C})$ (cf. T. XIV. 4.4 que da una demostración más rápida de d)).

*31. Sean $a_1, a_2, \dots, a_n \in \mathbf{C}$ tales que a_1, a_2, \dots, a_{n-1} no sean todos nulos. Determinar los casos en que es diagonalizable la matriz

$$A = \begin{bmatrix} 0 & \dots & 0 & a_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & . \\ a_1 & \dots & \dots & a_n \end{bmatrix}.$$

*32. Probar que existe una función $\varphi(n)$ con valores enteros, definida para el entero $n \geq 1$, con la siguiente propiedad:

Si $A \in \text{GL}(n, \mathbf{C})$ es una matriz con coeficientes en \mathbf{Z} , cuyo orden en el grupo $\text{GL}(n, \mathbf{C})$ (cf. def. II.4.4) es finito e igual a $\omega(A)$, se tiene:

$$\omega(A) \leq \varphi(n)$$

(acotar superiormente los coeficientes del polinomio característico).

*33. Toda matriz cuadrada $A \in \mathcal{M}_n(\mathbf{C})$ es semejante a su transpuesta. (Utilizar la reducción de Jordan.)

*34. (Otra demostración del teorema de Hamilton-Cayley para $M_n(\mathbf{C})$.)

Sea $X = [x_{ij}]_{1 \leq i \leq n, 1 \leq j \leq n}$ un elemento de $\mathcal{M}_n(\mathbf{C})$.

a) Demostrar directamente que, si X es diagonalizable, se tiene $P_X(X) = 0$.

b) Sea $\Delta((x_{ij}))$ el discriminante del polinomio $P_X(\lambda)$, (con una variable λ).

Probar que el polinomio $\Delta((X_{ij}))$ en las n^2 variables X_{ij} es no nulo.

c) Aplicando el teorema XIV.4.4. a los polinomios $P_X(X)$ y $\Delta(X)$, en las n^2 variables X_{ij} , demostrar que se verifica $P_X(X) = 0$ para toda matriz $X \in \mathcal{M}_n(\mathbf{C})$.

*35. Utilizar un método análogo al del ejercicio precedente para hallar que las matrices AX y XA (en donde $A, X \in \mathcal{M}_n(\mathbf{C})$) tienen el mismo polinomio característico.

(Empezar por el caso en que X es invertible.)

CAPÍTULOS XII Y XIII

1. Reducir a suma de cuadrados independientes las formas que siguen:

$$9x^2 - 6y^2 - 8z^2 + 6xy - 14xz + 18xw + 8yz + 12yw - 4zw, \\ x_1^2 + x_2^2 + x_3^2 - 2x_4^2 - 2x_1x_2 + 2x_1x_3 - 2x_1x_4 + 2x_2x_3 - 4x_2x_4.$$

2. Sea F_1 la forma cuadrática con n variables:

$$F_1 = \sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j.$$

Probar que existen formas lineales independientes (y_i) tales que

$$F_1 = y_1^2 + \frac{3}{4}y_2^2 + \frac{4}{6}y_3^2 + \cdots + \frac{n+1}{2n}y_n^2.$$

Precisar las y_i .

3. Si $A \in GL(n, \mathbf{R})$, la matriz $B = {}^t A \cdot A$ es definida positiva. Inversamente, si B es una matriz simétrica definida positiva de orden n , existe $A \in GL(n, \mathbf{R})$ tal que $B = {}^t A \cdot A$.

4. Si $A \in \mathcal{M}_n(\mathbf{C})$ es antisimétrica, $\exp(A) = \sum_{n \geq 0} \frac{A^n}{n!}$ es ortogonal. Asimismo, si $A + {}^t \overline{A} = 0$, $\exp(A)$ es unitaria (cf. ejercicio IX.14).

(Si A no es nilpotente, remitirse a la definición de $\exp(A)$ dada en el Curso de Análisis.)

5. Sean $A \in GL(n, \mathbf{R})$, $A = [a_{ij}]$. Probar que $[\det(A)]^2 \leq \prod_{1 \leq j \leq n} \left(\sum_{i=1}^n a_{ij}^2 \right)$. (Desigualdad de Hadamard.)

(Razonar por recurrencia, eligiendo una base conveniente.)

6. Sea A una matriz cuadrada de orden n con coeficientes reales, y simétrica. Se supone que existe un entero $k \geq 2$ tal que $A^k = I_n$ (matriz unidad de orden n). Probar que $A^2 = I_n$.

7. Probar que $M = \frac{1}{3} \begin{bmatrix} -2 & -1 & 2 \\ 2 & -2 & 1 \\ 1 & +2 & 2 \end{bmatrix}$ es una matriz de rotación. Determinar su ángulo y el eje.

(Ecole polytechnique.)

8. Sea $A \in O(n, \mathbf{R})$. Probar que si -1 no es valor propio de A , existe una matriz $Q \in \mathcal{M}_n(\mathbf{R})$, antisimétrica tal que

$$A = (I_n + Q)^{-1} (I_n - Q) = (I_n - Q) (I_n + Q)^{-1},$$

y que se verifica $A \in SO(n, \mathbf{R})$. ¿Recíproco?

(Representación paramétrica de Cayley de $SO(n, \mathbf{R})$.)

9. Sea E un espacio euclídeo de dimensión p . A cada n -pla (x_1, x_2, \dots, x_n) de elementos de E se le asocia el número

$$G(x_1, \dots, x_n) = \det ([x_i \cdot x_j]_{i,j=1,2,\dots,n})$$

(determinante de Gram).

a) Probar que x_1, x_2, \dots, x_n están ligados si, y sólo si, $G(x_1, \dots, x_n) = 0$; y probar que si x_1, \dots, x_n son independientes, se tiene $G(x_1, \dots, x_n) > 0$.

b) Probar que, para toda permutación $\sigma \in \mathfrak{S}_n$, se tiene $G(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = G(x_1, \dots, x_n)$, y que el valor de $G(x_1, \dots, x_n)$ no se modifica si se añade a uno de los vectores, por ejemplo a x_i , una combinación lineal de los restantes vectores $x_j (j \neq i)$. Calcular $G(\lambda x_1, x_2, \dots, x_n)$.

c) Suponemos que x_1, \dots, x_n son independientes. Sea $x \in E$, y sea $d(x, H)$ la distancia de x al hiperplano $H = \text{Vect}(x_1, \dots, x_n)$. Establecer la fórmula

$$d(x, H)^2 = \frac{G(x, x_1, \dots, x_n)}{G(x_1, \dots, x_n)}.$$

10. En el espacio vectorial E se llama norma a toda aplicación N de E en \mathbf{R}^+ que verifique las condiciones 1, 2, 3 del teorema XII.5.3.

Para que una norma N en \mathbf{R}^n sea euclídea (e.d. asociada a una forma cuadrática definida positiva) es necesario y suficiente que, para todo $x \in \mathbf{R}^n$ y todo $y \in \mathbf{R}^n$, se verifique (teorema de la mediana):

$$[N(x+y)]^2 + [N(x-y)]^2 = 2[N(x)]^2 + 2[N(y)]^2.$$

(Con la ayuda de N se definirá el producto escalar.)

11. Sea (e_1, \dots, e_n) una base ortonormal del espacio euclídeo E . Sea $E_p = \text{Vect}(e_1, \dots, e_p)$ y sea \vec{u} el vector unitario de componentes $\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)$.

¿Qué ángulo forman \vec{u} y E_p (Utilizar el ejercicio 9.)

12. Ecuaciones reducidas, en coordenadas rectangulares, de las siguientes cuádricas de \mathbf{R}^3 :

$$(x+y)(y-z) + 3x - 5y = 0, \\ x^2 - 2y^2 - z^2 - 4yz + 2xz + 2y + 2z + 3 = 0.$$

13. a) Con una transformación ortogonal real, reducir las formas que siguen a sumas de cuadrados:

$$F_1 = \sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j, \quad F_2 = \sum_{1 \leq i < j \leq n} x_i x_j.$$

b) Utilizar la matriz $[(\omega_k)^j]_{\substack{0 \leq k \leq n-1 \\ 0 \leq j \leq n-1}}$, en donde $\omega_k = e^{2ik\pi/n}$, para reducir F_1 y F_2 a sumas de cuadrados por medio de transformaciones unitarias.

14. En \mathbb{R}^n euclídeo, se consideran las cuádricas de ecuaciones: $B_1(x) = 0$, $B_2(x) = 0$, en donde las matrices de B_1 y de B_2 son, respectivamente, ${}^t A \cdot A$ y $A \cdot {}^t A$ (A designa una matriz cuadrada cualquiera). Probar que estas cuádricas son isométricas. (Empezar considerando el caso en que A es invertible).

15. Base ortonormal de vectores propios para la matriz de orden n :

$$M = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 1 \\ \vdots & & & 1 & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}.$$

*16. En el espacio proyectivo $\mathcal{P}_2(\mathbb{C})$ se consideran dos cónicas propias distintas Γ y Δ . En una referencia proyectiva cualquiera, las matrices de las ecuaciones de Γ y de Δ se designan por A y B . Demostrar que $\text{Tr}(A^{-1}B) = 0$ es la condición necesaria y suficiente para que existan tres puntos de Δ que constituyan un triángulo autopolar respecto de Γ .

17. Se designa por (e_{ij}) ($i, j = 1, 2, \dots, n$) la base canónica de $\mathcal{M}_n(K)$; y, para toda matriz $A = \sum_{i,j} a_{ij} e_{ij}$ de $\mathcal{M}_n(K)$, se designa por $\text{Tr}(A) = \sum_i a_{ii}$ a la traza de A .

a) Probar que la aplicación $T: (A, B) \rightarrow \text{Tr}(AB)$ es una forma bilineal simétrica no degenerada definida en $\mathcal{M}_n(K)$.

b) Probar que las matrices $f_{ij} = 1/2 (e_{ij} + e_{ji})$ y $g_{ij} = 1/2 (e_{ij} - e_{ji})$ ($i, j = 1, 2, \dots, n$) constituyen una base ortonormal respecto de T .

(MP2, París.)

18. Se dan dos formas hermíticas positivas [resp. definidas positivas], referidas a una misma base

$$\varphi(x) = \sum_{i,j} a_{ij} x_i \bar{x}_j, \quad \psi(x) = \sum_{i,j} b_{ij} x_i \bar{x}_j.$$

a) Probar que la forma

$$R(x) = \sum_{i,j} a_{ij} b_{ij} x_i \bar{x}_j$$

es hermítica positiva [resp. definida positiva]. (Estudiar en primer lugar el caso en que ψ posee rango 1.)

b) La misma cuestión para

$$E(x) = \sum_{i,j} e^{a_{ij}} x_i \bar{x}_j.$$

(Ecole polytechnique.)

c) Si φ es definida positiva, y ψ positiva no nula, se tiene $\sum_{i,j} a_{ij} b_{ij} > 0$.

19. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica cualquiera. Demostrar que

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

es simétrica, y definida positiva.

(Utilizar una base de vectores propios de A .)

20. a) Sean A una matriz simétrica positiva y k un entero > 1 . Probar que todo vector propio de A^k es un vector propio de A . (Considerar una base de vectores propios de A^k).

b) Sean A, B dos matrices simétricas positivas. Probar que para todo entero $k > 1$, la relación $A^k = B^k$ implica $A = B$.

*21. Sea $A \in \mathcal{M}_n(\mathbf{R})$ una matriz simétrica definida positiva, y sea (e_1, \dots, e_n) una base de vectores propios de A , siendo e_i el vector propio asociado al valor propio $\lambda_i (i = 1, 2, \dots, n)$. Designamos por P a la matriz de cambio de base que pasa de la base canónica de \mathbf{R}^n a la base (e_i) , y se define

$$\text{Log}(A) = PAP^{-1}, \text{ en donde } A = \begin{bmatrix} \text{Log } \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & \dots & 0 & \text{Log } \lambda_n \end{bmatrix}.$$

a) Probar que esta definición es independiente de la elección de la base (e_i) . ¿Es simétrica la matriz $\text{Log } A$?

b) Probar que $\exp(\text{Log } A) = A$. Recíprocamente, si B es simétrica, $\exp(B)$ es simétrica definida positiva (ejercicio XII.19). Probar, que entonces, $\log(\exp(B)) = B$.

Calcular $\log(A_1 A_2)$ cuando $A_1 A_2 = A_2 A_1$, en donde A_1 y A_2 son simétricas definidas positivas.

c) Sea $M \in \text{GL}(n, \mathbf{R})$. La matriz $A = {}^t M M$ es definida positiva. Se pone

$$Q = \exp\left(\frac{1}{2} \log A\right).$$

Probar que $H = MQ^{-1}$ es ortogonal.

d) Probar que la descomposición $M = HQ$ (H ortogonal, Q simétrica definida positiva) es única. (Utilizar el resultado b) del ejercicio XII.20, con $k = 2$.)

*22. Sean E un espacio hermitico de dimensión finita, y u un endomorfismo unitario de E . Se define $v = e - u$ (en donde e es la aplicación idéntica de E). Probar que el núcleo N de v es el ortogonal de la imagen de v , y que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{p=0}^{n-1} u^p(x)$$

es la proyección ortogonal de x sobre N .

(Ecole polytechnique.)

*23. Sea G un subgrupo finito de $\text{GL}(n, \mathbf{R})$ [resp. $\text{GL}(n, \mathbf{C})$]. Probar que existe una forma cuadrática q definida positiva en \mathbf{R}^n [resp. una forma hermitica h definida positiva en \mathbf{C}^n] invariante para todo $\varphi \in G$, e.d. tal que

$$\forall \varphi \in G, \quad \forall x \in \mathbf{R}^n, \quad q(\varphi(x)) = q(x) \quad [\text{resp. } \forall \varphi \in G, \quad \forall x \in \mathbf{C}^n, \quad h(\varphi(x)) = h(x)].$$

(Para cada elemento φ de G construir una forma cuadrática [resp. hermitica] invariante por φ .)

Problemas

P1

Los polinomios considerados en este problema poseen coeficientes reales.

I. 1) Si α, β, γ son tres números reales distintos dos a dos, probar que

$$A(x) = \frac{(x - \beta)(x - \gamma)}{(\alpha - \beta)(\alpha - \gamma)}, \quad B(x) = \frac{(x - \gamma)(x - \alpha)}{(\beta - \gamma)(\beta - \alpha)}, \quad C(x) = \frac{(x - \alpha)(x - \beta)}{(\gamma - \alpha)(\gamma - \beta)}$$

forman una base en el espacio vectorial X de los polinomios de grado dos, a lo sumo.

2) Descomponer $1, x, x^2$ en esta base.

3) En lo que resta de I diremos que dos polinomios P y Q de $\mathbf{R}[x]$ son congruentes y escribiremos $P \simeq Q$ si $P - Q$ es un múltiplo de $(x - \alpha)(x - \beta)(x - \gamma)$. Designaremos por \tilde{P} la clase de P .

a) Probar que a cada polinomio P le corresponde un único polinomio \bar{P} de \tilde{P} perteneciente a X .

b) Dar la expresión de \bar{P} en la base A, B, C .

c) Probar que la aplicación $u: \bar{P} \mapsto \bar{x} \cdot \bar{P}$ es un endomorfismo de X . ¿Constituye un automorfismo de X ?

¿Qué se puede decir de $\text{Ker}(u - \alpha)$, $\text{Ker}(u - \beta)$, $\text{Ker}(u - \gamma)$?

Diagonalizar el endomorfismo v de X definido por $\bar{P} \mapsto \bar{x}^k \cdot \bar{P}$ ($k \in \mathbf{N}^*$).

II. Generalización. Utilización de las funciones simétricas de las raíces de un polinomio $A(x)$.

Sean $K = \{1, 2, \dots, n\}$ (n entero positivo), y $\alpha_1, \alpha_2, \dots, \alpha_n$ n números reales y distintos. Ponemos $A(x) = \prod_{i \in K} (x - \alpha_i)$, y, para todo $i \in K$, $A_i(x) = \prod_{j \in K - \{i\}} (x - \alpha_j)$: por lo tanto, A_i es A desprovisto de su factor $x - \alpha_i$, y escribimos $a_i = 1/A_i(\alpha_i)$.

Finalmente, si C_k designa, para cada $k \in K$, el conjunto de las combinaciones de los n objetos de K tomados de k en k , ponemos:

$$\sigma_0 = 1, \quad \sigma_k = \sum_{S \in C_k} \prod_{i \in S} \alpha_i$$

en donde $\sigma_{k,p}$ designa el valor que toma la expresión precedente al substituir en ella σ_p por 0.

1) Sea M la matriz cuadrada de orden n , de término general

$$m_{k,p} = (-1)^{n+k} a_p \sigma_{n-k,p}.$$

Probar que M es invertible. Calcular M^{-1} y el determinante $\det(M)$ de M .

2) Sea Δ la matriz diagonal cuadrada de orden n cuyos elementos diagonales son $\delta_{ii} = \alpha_i$ para cada $i \in K$. Calcular $M^{-1}\Delta M$.

3) Se pone $B_k(x) = a_k \cdot A_k(x)$, $k \in K$.

Sea X el espacio vectorial de los polinomios de grado $(n-1)$ como máximo.

Decimos ahora que dos polinomios P y Q son congruentes y escribimos $P \simeq Q$ si $P - Q$ es un múltiplo de A .

Dado el polinomio P , \tilde{P} y \bar{P} designan, respectivamente, la clase de P y el representante de esta clase que pertenece a X . Finalmente, sea u el endomorfismo de X :

$$\bar{P} \mapsto \bar{x} \cdot \bar{P}.$$

a) Calcular $B_k(u) \circ B_{k'}(u)$, $(k, k') \in K^2$.

b) Calcular $\sum_{k \in K} B_k(u)$.

c) Calcular $\sum_{k \in K} \alpha_k B_k(u)$.

d) Si S es un polinomio dado de $\mathbf{R}[X]$, probar que la aplicación $\bar{P} \mapsto \bar{S} \cdot \bar{P}$ es un endomorfismo de X .

Probar que este endomorfismo es diagonalizable y diagonalizarlo.

Caracterizar \bar{S} para que éste sea un automorfismo.

(MP2, París, 1970).

P2

Para todo sistema (a_1, \dots, a_m) de números reales, se designa por $\mathbf{Q}[a_1, a_2, \dots, a_m]$ al menor subanillo de \mathbf{R} que contiene a \mathbf{Q} y a_1, a_2, \dots, a_m , que es precisamente el conjunto de los números reales de la forma $x = P(a_1, a_2, \dots, a_m)$, en donde P es un polinomio arbitrario con m variables y coeficientes en \mathbf{Q} .

I. 1) Probar que, si a_1, a_2, \dots, a_m son enteros, $\mathbf{Q}[\sqrt{a_1}, \sqrt{a_2}, \dots, \sqrt{a_m}]$ es un cuerpo. Dar en función de m una cota superior de la dimensión del \mathbf{Q} -espacio vectorial

$$\mathbf{Q}[\sqrt{a_1}, \sqrt{a_2}, \dots, \sqrt{a_m}].$$

2) Se llama *entero carente de cuadrado* a todo entero r que se escriba $r = p_1 p_2 \dots p_k$, en donde los p_i son números primos distintos dos a dos. Los números carentes de cuadrado r_1, r_2, \dots, r_n son *independientes* si se verifica la condición siguiente: para todo entero $i \in \mathbf{N}_n^*$, existe un número primo p_i que divide a r_j y no a ningún r_i para $j \neq i$. Para toda parte H de \mathbf{N}_n^* se define $t_H = \prod_{i \in H} r_i$, y por convenio, si H es vacío, $t_H = 1$.

Se llama \mathcal{F}_n a la siguiente propiedad: «Para todo sistema de n enteros carentes de cuadrado independientes r_1, r_2, \dots, r_n , los elementos $\sqrt{t_H}$, en donde H recorre el conjunto de partes de \mathbf{N}_n^* , constituyen una base del \mathbf{Q} -espacio vectorial $\mathbf{Q}[\sqrt{r_1}, \sqrt{r_2}, \dots, \sqrt{r_n}]$ ». Demostrar que \mathcal{F}_n implica \mathcal{F}_{n+1} . A este fin, si r_1, r_2, \dots, r_{n+1} son números carentes de cuadrado e independientes, se deberá demostrar que

$$r_{n+1} \notin \mathbf{Q}[\sqrt{r_1}, \sqrt{r_2}, \dots, \sqrt{r_n}].$$

Para terminar, se utilizará una relación de la forma

$$\sqrt{r_{n+1}} = \sum_{H \in \mathbf{N}_n^*} \lambda_H \sqrt{t_H} \quad (\lambda_H \in \mathbf{Q} \text{ para todo } H).$$

3) Se conservan las hipótesis de 2). Si $x \in \mathbf{Q}[\sqrt{r_1}, \sqrt{r_2}, \dots, \sqrt{r_n}]$, describir un método para calcular $1/x$ como combinación lineal de los t_H ($H \in \mathbf{N}_n^*$) con coeficientes en \mathbf{Q} .

Aplicación: poner $1/x$ en la forma

$$A\sqrt{2} + B\sqrt{3} + C\sqrt{5} + D\sqrt{6} + E\sqrt{10} + F\sqrt{15} + G\sqrt{30}$$

cuando $x = \sqrt{2} + \sqrt{5} + \sqrt{6} + \sqrt{15}$.

II. 1) Se hace $\alpha = \sqrt[3]{2}$. Por medio de la relación $\alpha^3 = 2$, probar que $\mathbf{Q}[\alpha]$ es un espacio vectorial de dimensión 3 sobre \mathbf{Q} y $(1, \alpha, \alpha^2)$ es una base.

Sean $x \in \mathbf{Q}[\alpha]$, $x \neq 0$. ¿Qué podemos decir de la aplicación \mathbf{Q} -lineal $\varphi_x : \mathbf{Q}[\alpha] \rightarrow \mathbf{Q}[\alpha]$ tal que $\varphi_x(y) = xy$ para $y \in \mathbf{Q}[\alpha]$? Deducir que $\mathbf{Q}[\alpha]$ es un cuerpo.

2) Probar que el polinomio $X^3 - 2$ es irreducible en $\mathbf{Q}[X]$. Si $a \in \mathbf{Q}$ y $b \in \mathbf{Q}$, deducir que los polinomios $X^3 - 2$ y $X^3 + aX + b$ son primos entre sí. Escribir $\frac{1}{\alpha^2 + a\alpha + b}$ en la forma

$A\alpha^2 + B\alpha + C$ ($A, B, C \in \mathbf{Q}$). (Utilícese el teorema de Bezout.) Aplicación: calcular $\frac{1}{\sqrt[3]{4} + 2\sqrt[3]{2} + 3}$

3) Se escribe $\alpha = \sqrt[3]{2}$ y $\beta = \sqrt[3]{3}$. Probar que $\mathbf{Q}[\alpha, \beta]$ es un cuerpo.

a) Se propone demostrar que $\beta \notin \mathbf{Q}[\alpha]$. Para ello, se razonará por reducción al absurdo del modo siguiente. La hipótesis $\beta \in \mathbf{Q}[\alpha]$ implica $\mathbf{Q}[\alpha] = \mathbf{Q}[\alpha, \beta] = K$, y $(1, \alpha, \alpha^2)$ y $(1, \beta, \beta^2)$ son dos bases distintas del \mathbf{Q} -espacio vectorial K . Haciendo $\beta = a + b\alpha + c\alpha^2$ ($a, b, c \in \mathbf{Q}$) se calculará el determinante y la traza del endomorfismo $x \mapsto \beta x$ del \mathbf{Q} -espacio vectorial K en cada una de las bases $(1, \alpha, \alpha^2)$ y $(1, \beta, \beta^2)$ y se llegará a una contradicción.

b) ¿Cuál es la dimensión del \mathbf{Q} -espacio vectorial $\mathbf{Q}[\alpha, \beta]$? Dar una base formada por ciertos productos $\alpha^p \beta^q$. Expresar en esta base el número real

$$(\sqrt[3]{2} + \sqrt[3]{3} + \sqrt[3]{4})^{-1}.$$

P3

I. Sea \mathcal{G} el grupo de las sustituciones pares del conjunto $\mathcal{O} = \{0, 1, 2, 3\}$ y sea $\sigma \in \mathcal{G}$ definida por $\sigma(0) = 0$, $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$.

1) El conjunto \mathcal{K} de los $\gamma \in \mathcal{G}$ tales que $\gamma(0) = 0$, ¿constituye un subgrupo de \mathcal{G} ? ¿Cuántos elementos tiene?

2) Probar que, para todo $i \in \mathcal{O}$, existe un único elemento τ_i de \mathcal{G} tal que $\tau_i(0) = i$ y tal que $\tau_i \circ \tau_i$ sea la identidad.

3) El conjunto \mathcal{H} de los elementos de \mathcal{G} cuyo cuadrado es la identidad ¿es un subgrupo de \mathcal{G} ? ¿Cuántos elementos tiene?

4) Probar que, para todo $\gamma \in \mathcal{G}$, existe un único $i \in \mathcal{O}$ tal que $\gamma \circ \tau_i$ pertenece a \mathcal{K} y un único $j \in \mathcal{O}$ tal que $\tau_j \circ \gamma$ pertenece a \mathcal{K} .

5) Calcular $\sigma \circ \tau_1 \circ \sigma^2$ y $\sigma^2 \circ \tau_1 \circ \sigma$. Sea \mathcal{G}' un subgrupo de \mathcal{G} al que pertenezcan σ y τ_1 . Comparar \mathcal{G} y \mathcal{G}' .

II. \mathbf{C} designa el cuerpo de los números complejos y \mathcal{M} el conjunto de las matrices invertibles $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con coeficientes complejos.

Sea \mathcal{P} un conjunto y sea ∞ un elemento de \mathcal{P} tal, que el complementario de $\{\infty\}$ en \mathcal{P} sea igual a \mathbf{C} .

1) Demostrar que, para todo $M \in \mathcal{M}$, existe una biyección única $q(M)$ y $\mathcal{P} \rightarrow \mathcal{P}$ tal que, si $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

si $z \in \mathbf{C}$ y si $cz + d \neq 0$, se tiene: $q(M)(z) = \frac{az + b}{cz + d}$;

si $z \in \mathbf{C}$ y si $cz + d = 0$, se tiene: $q(M)(z) = \infty$.

2) Para M y N pertenecientes a \mathcal{M} , calcular $q(MN)$.

3) Sea $I = \{0, 1, r, r^2\}$, en donde $r = e^{2\pi i/3}$. Sea G el grupo de las aplicaciones biyectivas $f: \mathcal{P} \rightarrow \mathcal{P}$ tales que:

$\alpha)$ Existe $M \in \mathcal{M}$ tal que $f = q(M)$.

$\beta)$ Para todo $z \in I$, se tiene $f(z) \in I$ y f induce una substitución par de I .

Sean f y g dos elementos de G tales que $f(z) = g(z)$ para todo $z \in I$. Probar que $f = g$.

4) ¿Cuál es el número de elementos de G ? (Considerar las matrices

$$S = \begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad T_1 = \begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix},$$

hacer $s = q(S)$, $t_1 = q(T_1)$ y utilizar la cuestión 5 de la primera parte).

5) Probar que la relación «existe $f \in G$ tal que $f(z) = z'$ » es una relación de equivalencia en \mathcal{P} ; sus clases de equivalencia se llaman órbitas de G .

6) Describir las órbitas de 0 y de ∞ .

7) Describir una órbita de G que tenga seis elementos.

III. Se considera la fracción racional $F(Z) = \frac{Z(Z^3 - 1)}{8Z^3 + 1}$.

1) ¿Cuáles son los ceros y los polos de F ?

2) Sea $\Phi: \mathcal{P} \rightarrow \mathcal{P}$ la aplicación definida por:

$$\alpha) \Phi(z) = F(z) \quad \text{si} \quad z \neq \infty \quad \text{y} \quad \text{si} \quad 8z^3 + 1 \neq 0.$$

$$\beta) \Phi(z) = \infty \quad \text{si} \quad z = \infty \quad \text{o} \quad \text{si} \quad 8z^3 + 1 = 0.$$

Probar que, para todo $f \in G$, existe un número complejo no nulo $c(f)$ tal que, para todo $z \in \mathcal{P}$, se verifica: $\Phi(f(z)) = c(f)\Phi(z)$. (Se conviene que $c(f) \cdot \infty = \infty$.)

3) Calcular $c(s)$ y $c(t_1)$. El conjunto H de los $f \in G$ tales que $c(f) = 1$, ¿es un subgrupo de G ? ¿Cuántos elementos tiene?

4) En \mathcal{P} se considera la relación de equivalencia «existe $h \in H$ tal que $h(z) = z'$ »; sus clases de equivalencia se llaman órbitas de H . Hallar las órbitas que no tienen cuatro elementos.

5) Probar que para que dos elementos z y z' de \mathcal{P} pertenezcan a la misma órbita de H , es necesario y suficiente que tengan la misma imagen por Φ .

6) Describir una aplicación $\Psi: \mathcal{P} \rightarrow \mathcal{P}$ tal que, para todo $u \in \mathcal{P}$, el conjunto $\Psi^{-1}(u)$ sea una órbita de G .

7) Sea $U(Z)$ una fracción racional con coeficientes complejos tal que verifique

$$U\left(\frac{aZ + b}{cZ + d}\right) = U(Z) \quad \text{para toda matriz} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M} \quad \text{tal que} \quad q\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) \in H.$$

Probar que existe una fracción racional con coeficientes complejos $V(Z)$ tal que $U(Z) = V(F(Z))$.
(Saint-Cloud, 1970.)

P4

1) Probar que el conjunto M de las matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con coeficientes reales tales que $ad - bc = 1$ forma un grupo multiplicativo cuyo elemento unidad se designará por e .

Sea $H = \{z \in \mathbb{C} \mid \operatorname{Im}(z) > 0\}$, sean $m \in M$ y $z \in H$, se escribe

$$mz = \frac{az + b}{cz + d}.$$

Probar que $mz \in H$.

Sean $m, m' \in M, z \in H$, probar que $m(m'z) = (mm')z$, $-m \in M$, $(-m)z = mz$, $ez = z$.

2) Probar que el subconjunto Γ de M tal que $a, b, c, d \in \mathbb{Z}$ es un subgrupo de M .

Se dota a Γ de la topología inducida por la de \mathbb{R}^4 . Probar que Γ es discreto.

Se escribe $s = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ y $t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Definir geoméricamente las transformaciones $z \rightarrow sz$ y $z \rightarrow tz$.

Probar que $\{-e, +e\}$ es un subgrupo normal de M y de Γ . Se escribirá $G = \Gamma / \{-e, +e\}$.

3) Sea $z_0 \in H$. Probar que el número de pares $(c, d) \in \mathbb{Z} \times \mathbb{Z}$ tales que $|cz_0 + d|^2 \leq 1$ es finito.

Calcular dicho número cuando $z_0 = \sqrt{3} + i/2$.

Deducir que la aplicación de Γ en \mathbb{R} , $g \rightarrow \operatorname{Im}(g(z_0))$ tiene un máximo cuando g recorre Γ , que alcanza dicho máximo y que si g_0 es un punto de Γ en que dicho máximo es alcanzado, $|g_0(z_0)| \geq 1$.

Sea

$$D = \left\{ z \in H \mid -\frac{1}{2} \leq R(z) \leq +\frac{1}{2}, |z| \geq 1 \right\}.$$

Probar que para todo $z \in H$ existe $g \in \Gamma$ tal que $gz \in D$.

Para $z = \sqrt{3} + \frac{i}{2}$, hallar g tal que $gz \in D$.

4) Sean $z \in H$ y $g \in G$. Se supone que uno de los puntos z y gz se halla en B y que el otro se halla en D . Probar que g es el elemento neutro de G . En un principio se podrá suponer que $\operatorname{Im}(z) \leq \operatorname{Im}(gz)$ y estudiar entonces los posibles valores de c y d correspondientes a g .

5) Sean $g \in D$, $g \in G$, $gz \in D$, $gz \neq z$. Probar que se verifica una de las tres posibilidades siguientes:

$$gz = tz, \quad gz = t^{-1}z, \quad gz = sz.$$

Precisar en cada caso la posición de z .

6) Sea $I(z) = \{g \in G \mid gz = z\}$. Probar que, para $z \in D$, $I(z)$ se reduce a un elemento salvo si $z = i$, $\rho = -\frac{1}{2} + \frac{i\sqrt{3}}{2}$, $-\bar{\rho}$. Precisar en estos casos la naturaleza de $I(z)$.

7) Probar que hay diez elementos de G tales que $g(D) \cap D \neq \emptyset$. Trazar las imágenes de D para estos diez elementos.

8) Sea G' el subgrupo de G engendrado por s y t .

Probar que todo $z' \in H$ se puede enviar a D por medio de un elemento de G' (se puede utilizar un razonamiento análogo al de 3)). Concluir que $G' = G$.

Mostrar que es posible obtener directamente la descomposición de $g \in G$ en función de s, t, t^{-1} en el ejemplo $g = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}$.

9) Sea $k \geq 2$ un entero. Se define

$$f_k(z) = \sum_{\substack{(m,n) \in \mathbf{Z} \times \mathbf{Z} \\ (m,n) \neq (0,0)}} \frac{1}{(mz + n)^{2k}}.$$

Probar que esta familia es sumable para todo $z \in H$.

Probar que es uniformemente sumable para $0 \leq R(z) \leq 1$; $\text{Im } z \geq \mu > 0$.

Probar que $f_k(z+1) = f_k(z)$.

Deducir que la familia es uniformemente sumable para $\text{Im } z \geq \mu > 0$.

Probar que $f_k\left(-\frac{1}{z}\right) = z^{2k} f_k(z)$ y que para $\gamma \in \Gamma$, $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$f_k\left(\frac{az + b}{cz + d}\right) = (cz + d)^{2k} f_k(z).$$

(Math. I, París, 1967.)

P5

(Aplicaciones geométricas de los cuaterniones.)

I. K designa el espacio vectorial \mathbf{R} , de dimensión 1 sobre \mathbf{R} , y E el espacio euclídeo \mathbf{R}^3 . Se dota a la suma directa $\Gamma = K \oplus E$ de su estructura de \mathbf{R} -espacio vectorial de dimensión 4, y de la ley de multiplicación siguiente:

Si $q_1 = (t_1, V_1)$ y $q_2 = (t_2, V_2)$, se define

$$q_1 q_2 = (t_1 t_2 - V_1 \cdot V_2, t_1 V_2 + t_2 V_1 + V_1 \wedge V_2).$$

1) Si $q = (t, V) \in \Gamma$, se define $\tilde{q} = (t, -V)$, $N(q) = q\tilde{q}$. ¿Cuáles son las propiedades de las aplicaciones $q \mapsto \tilde{q}$ y $q \mapsto N(q)$? Probar que Γ es un cuerpo no conmutativo (cuerpo de los cuaterniones).

2) Se pone $e = (1, 0, 0, 0)$, $i = (0, 1, 0, 0)$, $j = (0, 0, 1, 0)$, $k = (0, 0, 0, 1)$. Construir la tabla del grupo multiplicativo engendrado por e, i, j, k en $\Gamma^* = \Gamma \setminus \{0\}$. Determinar el centro de Γ^* . Determinar todos los cuerpos L tales que $K \subset L \subset \Gamma$.

3) Sea S^3 el conjunto de los $q \in \Gamma$ tales que $N(q) = 1$, que es un subgrupo del grupo multiplicativo Γ^* . Se designa por S^2 al conjunto de los $V = (x, y, z) \in E$ tales que $x^2 + y^2 + z^2 = 1$ y por \mathcal{R}_3 al grupo de los giros de E que dejan fijo el origen.

a) Todo $q \in S^3$ se puede escribir $q = (\cos \alpha, \sin \alpha V)$, en donde $V \in S^2$. A cada $q \in S^3$ se asocia la aplicación $q^* : \Gamma \rightarrow \Gamma$ definida por $q^*(r) = qrq^{-1}$ para $r \in \Gamma$. Probar que la restricción q^{**} de q^* a E es un elemento de \mathcal{R}_3 del que se precisará el eje y el ángulo.

Aplicación. Escribir la matriz del giro de ángulo 2α , y cuyo eje está definido por el vector unitario $V = (\lambda, \mu, \nu)$.

b) La aplicación $p_1 : q \rightarrow q^{**}$ de S^3 en \mathcal{R}_3 es un homomorfismo epiyectivo de grupos. ¿Cuál es su núcleo?

4) A cada $q = (t; x, y, z) \in \Gamma$, se asocia la matriz $\psi(q)$ de $\mathcal{M}_2(\mathbf{C})$:

$$\psi(q) = \begin{bmatrix} u & -\bar{v} \\ v & \bar{u} \end{bmatrix} \text{ donde } u = t + iz, \quad v = ix + y.$$

Probar que ψ es un isomorfismo de I en un subanillo de $\mathcal{M}_2(\mathbf{C})$ y que la restricción de ψ a S^3 es un isomorfismo de S^3 sobre el grupo multiplicativo $SU(2, \mathbf{C})$ de las matrices especiales unitarias (e.d. matrices A tales que ${}^t A \bar{A} = I$ y $\det(A) = 1$).

II. E se halla referido a su referencia ortonormal canónica $Oxyz$; N designa el punto $(0, 0, 1)$; las notaciones de I se conservan.

Se identifica el plano xOy con el cuerpo \mathbf{C} de los números complejos. Si $P \in \mathbf{C}$, la recta NP vuelve a cortar a S^2 en un punto $Q \neq N$. La aplicación $\sigma : P \rightarrow Q$ es una biyección de \mathbf{C} en $S^2 \setminus \{N\}$ (proyección estereográfica).

1) Calcular las coordenadas (X, Y, Z) de Q en función de z y \bar{z} , en donde z designa el afijo de P .

2) A cada matriz $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de $\mathcal{M}_2(\mathbf{C})$, de determinante no nulo, se le asocia la homografía $f_M : \tilde{\mathbf{C}} \rightarrow \tilde{\mathbf{C}}$ tal que

$$f_M(z) = \frac{az + b}{cz + d}$$

($\tilde{\mathbf{C}}$ es la esfera de Riemann obtenida añadiendo a \mathbf{C} un elemento ∞ tal que $\sigma(\infty) = N$.)

a) Comprobar que $M \rightarrow f_M$ es un homomorfismo de $GL(2, \mathbf{C})$ en el grupo de las transformaciones circulares. Precisar su imagen y su núcleo.

b) Probar que el grupo S^3 está engendrado por los elementos de la forma $(\cos \alpha, 0, \sin \alpha, 0)$ y $(\cos \beta, 0, 0, \sin \beta)$, α y β recorren \mathcal{R} independientemente.

c) Sea $q \in S^3$. Probar que la restricción de q^{**} a S^2 es la transmutada de $f_{\psi(q)}$ por σ , es decir, que, para todo $p \in \tilde{\mathbf{C}}$, se tiene: $\sigma^{-1} q^{**} \circ \sigma(P) = f_{\psi(q)}(P)$. ¿Conclusión?

P6

Preámbulo:

Demostrar las fórmulas:

$$S = \cos a + \cos(a + h) + \cdots + \cos[a + (n - 1)h] = \frac{\sin \frac{nh}{2} \times \cos \left[a + (n - 1) \frac{h}{2} \right]}{\sin \frac{h}{2}}$$

$$S' = \sin a + \sin(a + h) + \cdots + \sin[a + (n - 1)h] = \frac{\sin \frac{nh}{2} \times \sin \left[a + (n - 1) \frac{h}{2} \right]}{\sin \frac{h}{2}}.$$

Se empezará considerando la suma $(S + iS')$.

II. Se consideran las dos sumas:

$$x = \cos 3a + \cos 5a + \cos 7a + \cos 11a,$$

$$y = \cos a + \cos 9a + \cos 13a + \cos 15a,$$

en donde se ha puesto $\pi/17 = a$.

Calcular $(x + y)$ y xy . Deducir los valores de x y de y , por medio de radicales de raíces cuadradas. Establecer a continuación:

$$\begin{aligned} z &= \cos 3a + \cos 5a & t &= \cos 7a + \cos 11a, \\ u &= \cos a + \cos 13a & v &= \cos 9a + \cos 15a. \end{aligned}$$

Calcular zt y uv . Deducir los valores de z , t , u , v .

Deducir, finalmente, de lo que precede el valor del $\cos(\pi/17)$ por medio de radicales de raíces cuadradas.

II. Fijado de antemano un número n entero positivo, se escribe,

$$a_j = (-1)^j \cos \frac{j\pi}{2n+1}.$$

en donde j designa un entero relativo.

1) Comparar a_j , a_{j+2n+1} , a_{2n+1-j} .

Establecer las fórmulas:

$$(1) \quad \sum_{j=1}^{j=n} a_j = -\frac{1}{2}$$

$$(2) \quad 2a_j a_k = a_{j+k} + a_{j-k}$$

$$(3) \quad 2a_j^2 = 1 + a_{2j}.$$

2) Utilizar estas fórmulas para calcular $\cos \pi/5$. (No se pide que se obtenga un valor decimal aproximado.)

3) En toda esta cuestión se toma $n = 8$.

Partiendo de a_1 , (3) permite el cálculo de a_2 , y, repitiendo la operación, de a_4 , después de a_8 ; si continuáramos obtendríamos de nuevo la misma sucesión de valores.

Ello nos conduce a considerar las sumas $x_1 = a_1 + a_2 + a_4 + a_8$ y $x_2 = a_3 + a_5 + a_6 + a_7$.

Calcular $(x_1 + x_2)$ y $x_1 x_2$. La fórmula (3) permite obtener una relación (4) entre a_j y a_{4j} : escribirla. Si se conoce a_j , (4) permite calcular a_{4j} . ¿Qué se obtendría si volviéramos a empezar? Demostrar que de esta manera es posible partir el conjunto de los a_j en cuatro subconjuntos; dos de estos subconjuntos forman x_1 , los otros dos x_2 . Calcular las sumas de los elementos de cada uno de estos subconjuntos. Deducir, finalmente, $\cos(\pi/17)$. ¿Se observa alguna relación entre este método de cálculo y el de la primera parte?

III. Se conservan las notaciones de la segunda parte y se toma $n = 3$.

Calcular:

$$(a_1 + a_2 + a_3), (a_1 a_2 + a_2 a_3 + a_3 a_1), a_1 a_2 a_3.$$

Formar la ecuación de tercer grado que tiene las raíces a_1, a_2, a_3 . Demostrar que no posee ninguna raíz racional. ¿Qué se obtiene para $\cos(\pi/7)$?

IV. Se conservan las notaciones de la segunda parte. Formar la relación (5) que relaciona a_j y a_{3j} .

Se toma $n = 6$ y se pone:

$$x_1 = a_1 + a_3 + a_4, \quad x_2 = a_2 + a_5 + a_6.$$

Inspirándose en lo que precede, reducir el cálculo de los a_j a la resolución de ecuaciones algebraicas de grados 2 y 3 que se establecerán. (No se pide resolver dichas ecuaciones.)

V. Gauss ha demostrado que la construcción de polígonos regulares con un número primo de lados, de la forma $(2n + 1)$, se reduce a resolver ecuaciones algebraicas cuyos grados son los factores primos de n .

Comprobar que este teorema es verdadero para un polígono regular de 19 lados. Resultará ventajoso, haciendo $n = 9$, considerar las tres sumas:

$$(a_1 + a_7 + a_8), (a_4 + a_6 + a_9) \text{ y } (a_2 + a_3 + a_5).$$

Tratar la misma cuestión para un polígono regular de 11 lados. En este caso se aconseja escribir el cociente de

$$(x^{11} - 1) \text{ por } (x - 1).$$

a continuación transformar la ecuación de grado 10 que se obtiene igualando a 0 el cociente precedente, haciendo:

$$x + \frac{1}{x} = 2a.$$

No se pide dar una demostración general del teorema de Gauss.

VI. Si p y q son dos números primos entre sí, enteros y positivos, demostrar que los enteros $p, 2p, \dots, kp, \dots, qp$ son no congruentes módulo q , dos a dos (se recuerda que, si a y b designan dos enteros, « a es congruente a b módulo q » significa « $a - b$ es divisible por q »). Deducir, de lo anterior, la existencia de dos enteros h y m tales que:

$$hp + mq = 1$$

y por consiguiente tales que:

$$h \cdot \frac{2\pi}{q} - m \cdot \frac{2\pi}{p} = \frac{2\pi}{pq}.$$

Deducir que si se sabe construir los polígonos regulares de p y q lados inscritos en la circunferencia de radio 1, se sabrá construir también los polígonos regulares de pq lados inscritos en la circunferencia antedicha. Es además evidente que sabremos construir los polígonos regulares de $2n$ lados, toda vez que sepamos construir los de n lados. Admitiendo el teorema de Gauss enunciado en la parte quinta, dar la lista de los polígonos regulares, con menos de 50 lados, cuya construcción se reduce a la resolución de ecuaciones algebraicas cuyos grados son inferiores o iguales a 3.

(*Inst. Nat. Agronomique*, 1970.)

P7

I. Se designa por G al grupo multiplicativo formado por las matrices

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

en donde a, b, c, d son reales y verifican $ad - bc = 1$, por H el subgrupo formado por las matrices de la forma

$$h_s = \begin{pmatrix} \cos s & \operatorname{sen} s \\ -\operatorname{sen} s & \cos s \end{pmatrix}$$

en donde s es real, por K el subgrupo de las matrices

$$k_t = \begin{pmatrix} \operatorname{ch} t & \operatorname{sh} t \\ \operatorname{sh} t & \operatorname{ch} t \end{pmatrix},$$

en donde t es real, y por L el subgrupo formado por las matrices

$$l_u = \begin{pmatrix} e^u & 0 \\ 0 & e^{-u} \end{pmatrix},$$

en donde u es real. Se designa por P al conjunto de los números complejos $z = x + iy$ que verifican $y > 0$, y para todo elemento g de G , se designa por U_g a la aplicación de P en sí mismo definida por

$$U_g(z) = \frac{az + b}{cz + d}.$$

- 1) Probar que la aplicación $g \mapsto U_g$ es un homomorfismo y precisar su núcleo.
- 2) Determinar las trayectorias de los puntos de P por la acción de H, K, L . (Se llama trayectoria de un punto z por la acción de H , por ejemplo, al conjunto de los puntos $U_g(z)$ cuando g recorre H .) A este fin se formarán

$$\frac{U_g(z) - i}{U_g(z) + i} \quad \text{y} \quad \frac{U_g(z) - 1}{U_g(z) + 1}.$$

- 3) Determinar el estabilizador del punto i en G , es decir, el conjunto de los elementos g de G que verifican $U_g(i) = i$.
- 4) Probar que, para todo punto z de P , existe un par y sólo uno (k, l) , en donde $k \in K$ y $l \in L$, tal que $U_{lk}(i) = z$. Deducir que todo elemento g de G se puede escribir de una manera y una sola en la forma $g = lkh$, en donde $l \in L$, $k \in K$, $h \in H$.
- 5) Si escribimos $z = x + iy = U_{l_k t}(i)$, expresar x e y en función de t y u , y a continuación t y u en función de x e y .
- 6) Si escribimos $z = x + iy = U_{h_s}(\lambda i)$, en donde $0 < \lambda \leq 1$, expresar x e y en función de λ y s .

II. En esta parte se consideran transformaciones lineales del espacio \mathbf{R}^3 en sí mismo de la forma $x \mapsto Ax$, en donde $(Ax)_i = \sum_{j=1}^3 a_{ij}x_j$, $i = 1, 2, 3$, y se considera también la forma cuadrática $F(x) = x_1^2 + x_2^2 - x_3^2$. Se designa por Γ al conjunto de las transformaciones A que conservan F , es decir, que verifican $F(Ax) = F(x)$ para todo $x \in \mathbf{R}^3$.

- 1) Probar que Γ es un grupo multiplicativo, y que si A pertenece a Γ se tiene

$$(\det A)^2 = 1 \quad \text{y} \quad a_{33}^2 \geq 1.$$

- 2) Sea E el conjunto de los puntos x de \mathbf{R}^3 que verifican $F(x) < 0$ y $x_3 > 0$. Probar que para un elemento A de Γ se tiene $a_{33} \geq 1$ si, y sólo si, $A(E) \subset E$. Se observará que si x, y son dos puntos de \mathbf{R}^3 tales que $x \in E$ y $-y \in E$ existe un punto z del segmento que une x con y tal que $F(z) \geq 0$. Deducir que el conjunto G' de los elementos A de Γ que verifican $\det A = 1$ y $a_{33} \geq 1$ es un grupo multiplicativo.

3) Se designa por Q al conjunto de los puntos x de \mathbf{R}^3 que verifican $F(x) = -1$ y $x_3 > 0$, por H' al subgrupo de G' formado por los elementos de la forma

$$h'_s = \begin{pmatrix} \cos s & \operatorname{sen} s & 0 \\ -\operatorname{sen} s & \cos s & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

en donde s es real, por K' al subgrupo de G' formado por los elementos de la forma

$$k'_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \operatorname{ch} t & \operatorname{sh} t \\ 0 & \operatorname{sh} t & \operatorname{ch} t \end{pmatrix},$$

en donde t es real, y por L' al subgrupo de G' formado por los elementos de la forma

$$l'_u = \begin{pmatrix} \operatorname{ch} u & 0 & \operatorname{sh} u \\ 0 & 1 & 0 \\ \operatorname{sh} u & 0 & \operatorname{ch} u \end{pmatrix},$$

en donde u es real. Determinar las trayectorias de los puntos de Q frente a la acción de H' , K' , L' , así como el estabilizador del punto $(0, 0, 1)$ en G' .

4) Probar que todo elemento g' de G' se puede escribir, de manera única, en la forma $g' = l' k' h'$ en donde $l' \in L'$, $k' \in K'$, $h' \in H'$.

III. Se designa por V al espacio vectorial formado por las matrices $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, en donde $\alpha, \beta, \gamma, \delta$ son reales que verifican $\beta = \gamma$, y por Φ al isomorfismo de \mathbf{R}^3 en V definido por

$$\Phi(x_1, x_2, x_3) = \begin{pmatrix} x_3 - x_1 & x_2 \\ x_2 & x_3 + x_1 \end{pmatrix}.$$

1) Probar que para todo elemento g de G la transformación T_g definida por

$$T_g(M) = g \cdot M \cdot {}^t g$$

es un automorfismo de V , a continuación, que $\Phi^{-1} \cdot T_g \cdot \Phi$ es un elemento de I' , y finalmente, que la aplicación $g \mapsto \Phi^{-1} \cdot T_g \cdot \Phi$ es un homomorfismo de G en I' . Este último homomorfismo se designará por Ψ .

2) Determinar el núcleo de Ψ .

3) Determinar las imágenes de G , H , K , L por Ψ .

4) De lo que precede, deducir una biyección Ω de P en Q tal que para todo $g \in G$ se verifique $\Omega \circ U_g \circ \Omega^{-1} = \Psi(g)$.

5) Hallar una función $\varphi : P \rightarrow \mathbf{R}$ continua y positiva tal que se verifique:

$$\iint_{\Omega} \varphi(x, y) \, dx \, dy = \iint_{U_g(\Omega)} \varphi(x, y) \, dx \, dy$$

para todo dominio acotado Ω de P limitado por un número finito de arcos diferenciables de P , y para todo $g \in G$.

(E. N. S. (Ulm), 2.ª prueba, 1970.)

P8

En todo el problema, F_q designa un cuerpo finito con q elementos ($q \geq 2$).

Se admite que un cuerpo de este tipo es conmutativo.

Para todo entero $n > 0$, E_n designa el espacio vectorial $(F_q)^n$ (espacio vectorial de dimensión n sobre F_q).

El objeto del problema es el estudio en I de ciertas propiedades «geométricas» del espacio \bar{E}_n en I y, para ciertos valores de q , del «grupo ortogonal» de orden n sobre F_q en II.

Las partes I y II son independientes. En II, la cuestión 4 es independiente de las cuestiones 1, 2, 3.

I. Sea n un entero > 0 . Se designa por $M_n(F_q)$ al anillo de las matrices cuadradas de orden n , con elementos en F_q , y por $GL(n, F_q)$ al grupo multiplicativo de los elementos invertibles de $M_n(F_q)$.

1) ¿Cuál es el número de elementos de E_n ?

2) Sea p un entero $\leq n$; probar, por recurrencia sobre p , que el número de sucesiones ordenadas (v_1, v_2, \dots, v_p) de p vectores libres de E_n es:

$$\lambda_{n,p} = (q^n - 1)(q^n - q) \dots (q^n - q^{p-1}) = \prod_{k=0}^{p-1} (q^n - q^k).$$

Deducir el número γ_n de elementos de $GL(n, F_q)$.

3) Sea (v_1, v_2, \dots, v_p) una sucesión fija de p vectores libres de E_n .

Se designa por H al subespacio engendrado por (v_1, v_2, \dots, v_p) en E_n . Probar que dar otra base (w_1, w_2, \dots, w_p) de H equivale a dar una matriz $P \in GL(p, F_q)$. Deducir el número $g_{n,p}$ de subespacios de dimensión p de E_n .

4) a) Sea H un subespacio de dimensión p de E_n . Determinar el número $\mathcal{A}_{n,H}$ de automorfismos lineales de E_n que dejan H globalmente invariante. (Examinese la matriz de tal automorfismo en una base adaptada a una descomposición $E_n = H + K$, en donde K es un suplementario de H .) Deducir el número de suplementarios de H en E_n .

b) Sean H_1, H_2 dos subespacios de dimensión p de E_n . Probar que existe por lo menos un automorfismo U de E_n tal que $U(H_1) = H_2$. Determinar el número de tales automorfismos.

5) H designa en todo momento un subespacio de dimensión p de E_n . Sea m un entero tal que $p \leq m \leq n$, y r un entero tal que $0 \leq r \leq n - p$.

a) Hallar el número de subespacios L de dimensión m de E_n tales que $L \supset H$. (Considerar el espacio vectorial cociente E_n/H).

b) Probar que el número de subespacios vectoriales M de E_n , de dimensión r , tales que $M \cap H = \{0\}$ es:

$$\prod_{k=0}^{r-1} \left(\frac{q^n - q^{p+k}}{q^r - q^k} \right).$$

6) Sea H un subespacio de dimensión p de E_n . Se designa por K a un subespacio fijo de dimensión $k \leq p$, de H .

Calcular el número de subespacios L , de dimensión m , de E_n , tales que $L \cap H = K$, en donde m es un entero que verifica $m - k \leq n - p$.

II. En esta parte se hacen las siguientes hipótesis:

● q es impar (de forma que en F_q la división por 2 es posible).

● En F_q existe un elemento θ tal que $\theta^2 = -1$.

Se designa por G_2 el grupo multiplicativo formado por los elementos de F_q , distintos de cero, que son el cuadrado de un elemento de F_q . Se admitirá, si es preciso, que G_2 tiene $(q-1)/2$ elementos y que $(x \notin G_2 \text{ e } y \notin G_2 \Rightarrow xy \in G_2)$.

1) Probar que existen $(2q - 1)$ pares ordenados (x_1, x_2) , elementos de $(F_q)^2$, tales que $x_1^2 + x_2^2 = 0$ y que, si b pertenece a F_q y $b \neq 0$, existen $(q - 1)$ pares ordenados (x_1, x_2) tales que $x_1^2 + x_2^2 = b$. (Transformar la expresión $x_1^2 + x_2^2$ en un producto.)

2) Sea m un entero ≥ 1 . Se llama λ_m el número de sucesiones ordenadas $(x_1, x_2, \dots, x_{2m})$ de elementos de F_q que verifican:

$$x_1^2 + x_2^2 + \dots + x_{2m}^2 = 0.$$

Para todo $b \neq 0$, con b perteneciendo a F_q , se llama $\mu_{m,b}$ al número de sucesiones $(x_1, x_2, \dots, x_{2m})$ de elementos de F_q que verifican:

$$x_1^2 + x_2^2 + \dots + x_{2m}^2 = b.$$

Probar, por recurrencia sobre m , que $\lambda_m = q^{2m-1} + q^m - q^{m-1}$ y $\mu_{m,b} = q^{2m-1} - q^{m-1}$.

3) m designa en todo momento un entero ≥ 1 . Utilizar los resultados de 2) para demostrar:

a) Que el número ν_m de sucesiones $(x_1, x_2, \dots, x_{2m+1})$ de elementos de F_q tales que:

$$x_1^2 + x_2^2 + \dots + x_{2m+1}^2 = 0,$$

está dado por $\nu_m = q^{2m}$.

b) Que el número ρ_m de sucesiones $(x_1, x_2, \dots, x_{2m+1})$ de elementos de F_q tales que

$$x_1^2 + x_2^2 + \dots + x_{2m+1}^2 = 1,$$

está dado por $\rho_m = q^{2m} + q^m$.

4) Se considera el espacio vectorial E_k . A toda forma bilineal simétrica F definida en E_k se le asocia la «forma cuadrática» f definida por $f(x) = F(x, x)$. Para cada base $B = (e_1, \dots, e_k)$ de E_k se designa por $\Delta(f, B)$ al determinante de la matriz de orden k cuyo término general es $F(e_i, e_j)$.

a) Probar que, si F y G son formas bilineales simétricas que definen la misma forma cuadrática f , entonces $F = G$.

En las cuestiones b, c, d, e y f que siguen, f designa una forma cuadrática definida en E_k y F la forma bilineal asociada.

b) Probar que, si $\Delta(f, B) \neq 0$ para una base B , entonces $\Delta(f, B) \neq 0$ para toda base B (cuando esto ocurra, diremos que « f es no degenerada»). Probar que, si $\Delta(f, B)$ pertenece a G_2 para una base B , entonces $\Delta(f, B)$ pertenece a G_2 para toda base B (cuando esto ocurra, diremos que « f es de primera especie»).

c) Se supone $k \geq 2$ y f no degenerada. Entonces existe $\varepsilon \in E_k$ tal que $f(\varepsilon) \neq 0$. Sea H_ε la recta engendrada por ε y

$$H_\varepsilon^\perp = \{ x \in E_k; F(x, \varepsilon) = 0 \}.$$

Probar que E_k es la suma directa de H_ε y de H_ε^\perp y que la restricción de f a H_ε^\perp es no degenerada. Deducir la existencia de una base (e_1, e_2, \dots, e_k) de E_k que es ortogonal respecto de f (es decir, tal que $F(e_i, e_j) = 0$ para todo $i \neq j$).

d) Suponemos $k = 2$ y f de primera especie. Sea (e_1, e_2) una base de E_2 , ortogonal respecto de f . Decimos que $\alpha_1 = f(e_1)$ y $\alpha_2 = f(e_2)$. Probar que $(\alpha_1, \alpha_2) \in G_2$. Para $x = x_1 e_1 + x_2 e_2$ perteneciente a E_2 , escribir $f(x)$ en forma de producto. Deducir la existencia de un ε_1 de E_2 tal que $f(\varepsilon_1) = 1$. Probar que es posible hallar ε_2 tal que $f(\varepsilon_2) = 1$ y que $F(\varepsilon_1, \varepsilon_2) = 0$.

e) Se supone $k \geq 3$ y f de primera especie. Sea (e_1, e_2, \dots, e_k) una base de E_k , ortogonal para f . Se escribe, para $1 \leq i \leq k$, $\alpha_i = f(e_i)$. Probar que existen i y j distintos tales que $(\alpha_i, \alpha_j) \in G_2$. Concluir, inspirándose en d), que existe un ε perteneciente a E_k tal que $f(\varepsilon) = 1$. Probar entonces que la restricción de f a $H_\varepsilon^\perp = \{x \in E_k; F(x, \varepsilon) = 0\}$ es de primera especie.

f) De lo antecedente deducir el resultado siguiente: Si f es una forma cuadrática de primera especie definida en E_k y si $k \geq 2$, existe en E_k una base $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k)$ ortonormal respecto de f , es decir que verifica:

$$f(\varepsilon_i) = 1 \quad (1 \leq i \leq k) \quad \text{y} \quad \text{para } i \neq j \quad F(\varepsilon_i, \varepsilon_j) = 0.$$

5) El espacio E_n ($n \geq 2$) se dota de la forma cuadrática:

$$f(x_1, x_2, \dots, x_n) = x_1^2 + x_2^2 + \dots + x_n^2.$$

Con la ayuda de los resultados obtenidos en 2), 3) y 4), probar que el número ω_n de sucesiones ordenadas de n vectores $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ de E_n , ortonormales para f , se obtiene por medio de las fórmulas siguientes:

Si $n = 2m + 1, \quad m \geq 1,$

$$\omega_n = 2(q^2 - 1)(q^4 - 1) \dots (q^{2m} - 1) q^{m^2}.$$

Si $n = 2m, \quad m \geq 2,$

$$\omega_n = 2(q^2 - 1)(q^4 - 1) \dots (q^{2m-2} - 1) q^{m(m-1)} (q^m - 1)$$

$$\omega_2 = 2(q - 1).$$

III. Esbozar el estudio análogo al realizado en II cuando en F_q no existe ningún elemento θ tal que $\theta^2 = -1$.

(Partes I y II: E. N. S. E. T., 1969.)

P9

Todas las curvas que intervienen en el problema son curvas del *plano proyectivo* $P_2(\mathbf{C})$ sobre el cuerpo \mathbf{C} de los números complejos. El espacio afín E_2 de dimensión 2 sobre \mathbf{C} se identifica con un subconjunto de $P_2(\mathbf{C})$, por medio de un sistema de ejes Oxy de E_2 , de la forma conocida. Intervendrá igualmente la *recta proyectiva* $P_1(\mathbf{C})$, identificada con $\mathbf{C} \cup \{\infty\}$, obtenido añadiendo a \mathbf{C} un «punto del infinito», designado por ∞ .

Se recuerdan las definiciones y propiedades siguientes:

● Una curva algebraica es un conjunto de puntos de $P_2(\mathbf{C})$, definido por una ecuación $F(X, Y, T) = 0$, en donde F es un polinomio homogéneo.

● Se admitirá que dos polinomios homogéneos *irreducibles* del mismo grado definen la misma curva si, y sólo si, son proporcionales.

● Un punto $I_0(X_0, Y_0, T_0)$ de la curva Γ de ecuación $F(X, Y, T) = 0$ es *ordinario* si $F'_{X_0}, F'_{Y_0}, F'_{T_0}$ no son todas nulas, *singular* en caso contrario.

● El punto $I_0(X_0, Y_0, T_0)$ de la curva Γ es un *punto de inflexión* si es ordinario, y si, para toda representación paramétrica propia $\rho \mapsto I_0 + \rho H$ de la tangente T_{I_0} a Γ en el punto I_0 (H es un punto de T_{I_0} distinto de I_0), la multiplicidad de la raíz $\rho = 0$ en la ecuación de las intersecciones de Γ y T_{I_0} es ≥ 3 .

● Se admitirá que las dos nociones precedentes son invariantes frente al grupo de las homografías de $P_2(\mathbf{C})$.

I. Sea a un número real > 0 , y \mathcal{F} la cúbica cuya representación paramétrica es:

$$X = 3at, \quad Y = 3at^2, \quad T = 1 + t^3, \quad t \in \mathbf{C}$$

1) a) Construir el conjunto de los puntos reales de \mathcal{F} . ¿Cuál es la ecuación cartesiana de \mathcal{F} ? Dar la condición para que tres puntos de parámetros t_1, t_2, t_3 estén alineados.

b) Buscar el conjunto \mathcal{O} de los puntos de inflexión. Probar que los puntos de \mathcal{O} están alineados.

2) Sea \mathcal{G} el grupo de las homografías que dejan a \mathcal{F} globalmente invariante. Para cada elemento h de \mathcal{G} , $\sigma(h)$ designa la restricción de h a \mathcal{O} .

a) Probar que la aplicación $\sigma : h \mapsto \sigma(h)$ es un homomorfismo de \mathcal{G} en el grupo de las biyecciones de \mathcal{O} en \mathcal{O} . ¿Cuál es el núcleo de σ ?

b) Sea I un elemento de \mathcal{O} . A todo M de \mathcal{F} , $M \neq I$, le hacemos corresponder el tercer punto de intersección N de la recta IM con \mathcal{F} . Probar que la aplicación $M \mapsto N$, convenientemente prolongada en I , es una biyección de \mathcal{F} en \mathcal{F} , y que es la restricción a \mathcal{F} de una homografía. Deducir que σ es epiyectiva.

¿A qué grupo muy simple es isomorfo el \mathcal{G} ?

II. 1) Sea $F(X, Y, T)$ un polinomio homogéneo de grado ≥ 3 , $I_0(X_0, Y_0, T_0)$ un punto de la curva I de ecuación $F = 0$. Se escribe la fórmula de Taylor de F en I_0 :

$$F(X_0 + \lambda, Y_0 + \mu, T_0 + \nu) = F(X_0, Y_0, T_0) + \\ - \lambda F'_{X_0} + \mu F'_{Y_0} + \nu F'_{T_0} + \frac{1}{2} (\lambda F'_{X_0} + \mu F'_{Y_0} + \nu F'_{T_0})^{(2)} + \dots$$

Se supone que I_0 es ordinario, se designa C_0 a la cónica de ecuación:

$$(XF'_{X_0} + YF'_{Y_0} + TF'_{T_0})^{(2)} = 0.$$

Probar que I_0 es un punto de inflexión de I si, y sólo si, la cónica C_0 es degenerada. (Será útil demostrar y utilizar el hecho: $I_0 \in C_0$.)

2) Sea $m \in P_1(\mathbf{C})$; I_m designa la cúbica de ecuación $F_m(X, Y, T) = 0$, con:

$$F_m(X, Y, T) = X^3 + Y^3 + T^3 - 3mXYT \quad \text{si } m \in \mathbf{C} \\ F_\infty(X, Y, T) = XYT.$$

Φ designa el conjunto de todas estas cúbicas.

a) ¿Para qué valores de m , I_m posee puntos singulares? Probar que, entonces, se descompone en tres rectas, que se escribirán. Construir los puntos reales de I_2 .

b) Deducir de 1) que, en general, los puntos de inflexión de I_m son exactamente los puntos de intersección de I_m y $I_{-\frac{m^3-4}{3m^2}}$. Probar que estos puntos de inflexión son siempre los siguientes:

$$\begin{array}{lll} A_0(0, 1, -1) & B_0(-1, 0, 1) & C_0(1, -1, 0) \\ A_1(0, 1, \alpha) & B_1(\alpha, 0, 1) & C_1(1, \alpha, 0) \\ A_2(0, 1, \beta) & B_2(\beta, 0, 1) & C_2(1, \beta, 0) \end{array}$$

en donde α, β son las raíces de $x^2 - x + 1 = 0$ (notación que utilizaremos en adelante). El conjunto de estos nueve puntos se designará \mathcal{O} . Demostrar que toda cúbica que contiene a \mathcal{O} es una I_m .

c) Comprobar que toda recta que pasa por dos puntos de \mathcal{O} pasa por un tercero y que no existen cuatro puntos de \mathcal{O} alineados. Calcular el número de rectas que los alinean. Calcular el número de ellas que contienen un punto dado de \mathcal{O} .

Probar que las cúbicas descompuestas de Φ son cuatro, que se designarán por (Γ_{m_i}) $1 \leq i \leq 4$ y que son los únicos sistemas de tres rectas cuya reunión contiene a \mathcal{G} .

3) Sea \mathcal{G}_g el grupo de las homografías que dejan a \mathcal{G} globalmente invariante, y, para $m \in \mathbb{C}$, \mathcal{G}_m el subgrupo de \mathcal{G}_g formado por las homografías que dejan Γ_m globalmente invariante.

a) Sean $h \in \mathcal{G}_g$ y $m' \in P_1(\mathbb{C})$; h transforma a $\Gamma_{m'}$ en $\Gamma_{m''}$. Se escribe $m'' = h(m')$. Probar que h es una homografía de $P_1(\mathbb{C})$. Probar que la restricción h de h al conjunto $D = \{m_1, m_2, m_3, m_4\}$ es una biyección de D . Establecer que h es una permutación par (se recuerda que cuando la razón doble (x_1, x_2, x_3, x_4) sólo toma 2 valores, estos valores son α y β , y las permutaciones de los x_i que la dejan invariante son las permutaciones pares).

b) Suponiendo que \tilde{h} sea una de las permutaciones:

$$\begin{pmatrix} m_1 & m_2 & m_3 & m_4 \\ m_2 & m_1 & m_4 & m_3 \end{pmatrix} \quad \begin{pmatrix} m_1 & m_2 & m_3 & m_4 \\ m_3 & m_4 & m_1 & m_2 \end{pmatrix} \quad \begin{pmatrix} m_1 & m_2 & m_3 & m_4 \\ m_4 & m_3 & m_2 & m_1 \end{pmatrix}$$

calcular \tilde{h} . Probar que, si además, $h \in \mathcal{G}_m$, m es uno de los seis valores:

$$(1) \quad 1 \pm \sqrt{3}, \quad j(1 \pm \sqrt{3}), \quad j^2(1 \pm \sqrt{3}) \quad \left(\text{con } j = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \right).$$

Suponiendo a continuación que \tilde{h} es una de las ocho permutaciones pares distintas de la identidad y que dejan invariante uno de los m_i , calcular \tilde{h} y precisar sus puntos dobles. Probar que, si además $h \in \mathcal{G}_m$ y $m \notin D$, m es uno de los cuatro valores:

$$(2) \quad 0, \quad -2, \quad -2j, \quad -2j^2.$$

c) Supongamos que $m \notin D$ y que m no es ninguno de los valores (1) o (2). Si $h \in \mathcal{G}_m$, ¿qué podemos decir de \tilde{h} ? Concluir que en este caso el conjunto E de tres rectas $A: (A_0 A_1 A_2)$, $B: (B_0 B_1 B_2)$, $C: (C_0 C_1 C_2)$ es invariante por h . Sea $\tau(h)$ la permutación

$$\begin{pmatrix} A & B & C \\ h(A) & h(B) & h(C) \end{pmatrix}$$

de E . Probar que $h \mapsto \tau(h)$ es un homomorfismo de \mathcal{G}_m en el grupo de las biyecciones de E en E . Determinar con cuidado el núcleo de τ .

d) Se conservan las hipótesis y las notaciones de c). Sea $I \in \mathcal{G}$. A todo punto M de Γ_m , $M \neq I$, le hacemos corresponder el tercer punto de intersección N de IM con Γ_m . Probar que la aplicación $M \mapsto N$ (convenientemente prolongada en I) es una biyección de Γ_m en Γ_m y que es la restricción a Γ_m de una homografía. De todo ello deducir que τ es epiyectiva. ¿Cuál es el número de elementos de \mathcal{G}_m ?

e) Se supone que $m = a_i$ es uno de los valores (2). Demostrar que existe un mismo elemento m_i de D fijo para todas las \tilde{h} para $h \in \mathcal{G}_{a_i}$. Γ_{m_i} es entonces un conjunto E_i de tres rectas. Inspirándose en c) y d) definir un homomorfismo ν de \mathcal{G}_{a_i} en el grupo de las biyecciones de E_i en E_i . Probar que ν es epiyectivo, y hallar su núcleo. ¿Cuál es entonces el número de elementos de \mathcal{G}_{a_i} ? (En esta cuestión, sólo se tratará a fondo el caso $a_i = 0$.)

f) Inspirándose en los métodos anteriores, imaginar un método que permita calcular el número de elementos de \mathcal{G}_g , y el de \mathcal{G}_m cuando m es uno de los valores de (1).

N.B. Se puede demostrar que toda cúbica con un punto doble que no sea de retroceso se deduce de \mathcal{F} por una homografía, y que toda cúbica sin puntos dobles se deduce de una de las Γ_m por una

homografía. Con otras palabras, el estudio abordado en el problema es general.

(No se pide comprobar estas propiedades.)

(E. N. S. E. T., 1970.)

P10

Todos los polinomios que intervienen en este problema son polinomios con coeficientes complejos. Se designa por P el plano, conjunto de pares (a, b) de números complejos. A un conjunto Γ de puntos de P se le llama curva algebraica de P si existe un polinomio no constante $f(x, y)$ con dos variables x e y , tal que el conjunto de los puntos (a, b) de P que verifican la condición $f(a, b) = 0$ sea Γ . Si f es de grado mínimo en el conjunto de los polinomios que poseen esta propiedad, se dice que $f(x, y) = 0$ es una ecuación minimal de Γ .

I. Se supone que el polinomio $f(x, y)$ es irreducible, es decir no constante ni producto de dos polinomios no constantes. Nos proponemos demostrar que $f(x, y) = 0$ es una ecuación minimal de la curva algebraica Γ , conjunto de los puntos (a, b) tales que $f(a, b) = 0$. Se designa por E el conjunto de los polinomios no constantes, en las variables x e y , que son nulos en todo punto de Γ y cuyo grado en y es minimal. Se designa por $g(x, y)$ a un elemento de E .

1) Inspirándose en la división euclídea, probar que si $F(x, y)$ es un polinomio no constante nulo en todo punto de Γ , existen polinomios no nulos $P(x)$ y $q(x, y)$ tales que $P(x)F(x, y) = q(x, y)g(x, y)$.

2) Sea a un número complejo. Probar que si el producto $q(x, y)g(x, y)$ es divisible por $(x - a)^r$ (en donde r es un entero > 0) y si $q(x, y)$ no es divisible por $x - a$, entonces $g(x, y)$ es divisible por $(x - a)^r$ (remitirse al caso $a = 0$).

3) Se supone que $P^0(x)$ es un polinomio de grado minimal en el conjunto de los polinomios no nulos $P(x)$ tales que $P(x)f(x, y)$ es p múltiplo de $g(x, y)$. Se define

$$P^0(x) f(x, y) = q^0(x, y) g(x, y).$$

Probar que $g(x, y)$ es de la forma $P^0(x)h(x, y)$, en donde $h(x, y)$ es un polinomio. Deducir, utilizando la irreducibilidad de f , que uno de los polinomios, bien q^0 o bien h , es constante.

a) Probar que si q^0 es constante, $g(x, y)$ es de la forma $cP^0(x)f(x, y)$, en donde c es constante.

b) En el caso en que h sea constante, demostrar que $f(x, y)$ es independiente de y y de grado 1.

Deducir de a) y de b) que, en todos los casos, $g(x, y)$ es múltiplo de $f(x, y)$ y que $f(x, y)$ pertenece a E .

4) Sea $F(x, y)$ un polinomio nulo en todo punto de Γ . Deducir de los resultados precedentes que $F(x, y)$ es un múltiplo de $f(x, y)$. Probar que $f(x, y) = 0$ es una ecuación minimal de Γ .

5) Probar que un polinomio $F(x, y)$ tal que Γ es el conjunto de los puntos (a, b) que verifican $F(a, b) = 0$ es necesariamente de la forma $cf(x, y)^r$, en donde c es una constante no nula y en donde r es un entero ≥ 1 .

II. C designará el conjunto de las ternas de números complejos, imagen de P por la aplicación φ que al punto (a, b) le asocia el punto (a^2, ab, b^2) .

1) a) Comprobar que la imagen de φ es el conjunto de los puntos (c, d, e) tales que $ce - d^2 = 0$ y determinar la imagen recíproca de un punto de C por medio de la aplicación φ .

b) Determinar la imagen por φ de una recta que pase por el origen $(0, 0)$ de P .

c) Determinar la imagen por φ de una recta de P que no pase por el origen.

2) Sea $f(x, y)$ un polinomio. Probar que existen tres polinomios $F_0(X, Y, Z)$, $F_1(X, Z)$, $F_2(X, Z)$ con las variables X, Y, Z , tales que:

$$f(x, y) = F_0(x^2, xy, Y^2) + xF_1(x^2, y^2) + yF_2(x^2, y^2).$$

Los polinomios $F_0(x^2, xy, y^2)$, $F_1(x^2, y^2)$, ¿están unívocamente determinados?

En lo que sigue de II, se supone que el polinomio $f(x, y)$ es irreducible. Nos proponemos estudiar las relaciones entre la curva algebraica Γ de ecuación minimal $f(x, y) = 0$ y su imagen $\varphi(\Gamma)$.

3) Supondremos que Γ es simétrica respecto del origen $(0, 0)$ de P . Deducir de I la existencia de una constante a tal que:

$$f(-x, -y) = af(x, y).$$

¿Cuáles son los valores posibles de a ?

4) Supondremos que Γ es simétrica respecto del origen y que el grado de f es un número par $2p$. Probar que existe un polinomio $F(X, Y, Z)$ tal que $\varphi(\Gamma)$ es el conjunto de los puntos (c, d, e) de C tal que $F(c, d, e) = 0$. ¿Cuál es el grado minimal de los polinomios F que poseen esta propiedad?

5) Supondremos que Γ es simétrica respecto del origen y que el grado de f es un número impar $2p + 1$. Probar que existe un polinomio $F(X, Y, Z)$ tal que $\varphi(\Gamma)$ es el conjunto de los puntos (c, d, e) de C tales que $F(c, d, e) = 0$.

¿Cuál es el grado mínimo de los polinomios F que poseen esta propiedad? Probar que existen dos polinomios $G(X, Y, Z)$ y $H(X, Y, Z)$ de grado $p + 1$ tales que $\varphi(\Gamma)$ es el conjunto de los puntos (c, d, e) de C que verifican las condiciones $G(c, d, e) = H(c, d, e) = 0$.

6) Se supone que $f(x, y) = x^3 - y^3 - x$. Probar la irreducibilidad de este polinomio. Sea V el conjunto de los polinomios g nulos en todo punto de Γ tales que $g(x, y) = g(-x, -y)$. ¿Cuál es la dimensión del espacio vectorial formado por los polinomios de grado ≤ 4 pertenecientes a V ? Hallar un polinomio de V con el menor grado posible que no sea nulo en ninguno de los puntos del complementario de Γ en P .

7) Se supone que Γ no es simétrica respecto del origen de P . Utilizando la expresión de f obtenida en 2), dar una ecuación minimal de la curva simétrica de Γ respecto del origen. Probar que existe un subconjunto \mathcal{E} de Γ tal que la restricción de φ a $\Gamma - \mathcal{E}$ es biunívoca y que la aplicación de $\varphi(\Gamma) - \varphi(\mathcal{E})$ en $\Gamma - \mathcal{E}$ recíproca de φ viene dada por las fórmulas:

$$x = u(X, Y, Z), \quad y = v(X, Y, Z),$$

en donde u, v, w son fracciones racionales con coeficientes complejos en X, Y, Z . Explícitense u, v, w por medio de los polinomios F_0, F_1, F_2 , introducidos en 2). Probar que es posible elegir el conjunto \mathcal{E} finito. (Para ello se admitirá el siguiente resultado: si dos polinomios g y h con las variables x e y carecen de divisores comunes no constantes, el sistema de ecuaciones $g(x, y) = h(x, y) = 0$ sólo posee un número finito de soluciones.)

(E. N. S. rue d'Ulm, 1965, extracto.)

(ver Cap. XIV).

Bibliografía

Obras que completan a la presente:

1. J. M. ARNAUDIÈS, *Les cinq polyèdres réguliers et leurs groupes* (C.D.U., Paris).
2. E. ARTIN, *Algèbre géométrique*. G. Villars, 1962.
3. Z. I. BOREVITCH et I. R. CHAFAREVITCH, *Théorie des nombres*. G. Villars, 1967.
4. N. BOURBAKI, *Théorie des ensembles*. Actualités Sc. et Ind., n^{os} 1212-1243-1258 (Hermann).
5. N. BOURBAKI, *Algèbre*. Actualités Sc. et Ind., n^{os} 934-1032-1102 (Hermann).
6. C. CHEVALLEY, *Theory of Lie groups I*. Princeton University Press, 1946.
7. P. y M. L. DUBREIL, *Lecciones de álgebra moderna*, Editorial Reverté, S.A., Barcelona.
8. R. GODEMENT, *Cours d'algèbre*. Hermann, 1966.
9. G. H. HARDY and E. M. WRIGHT, *An introduction to the theory of numbers*. Oxford Clarendon Press, 1954.
10. S. MAC LANE et G. BIRKHOFF, *Algèbre*. Gauthier-Villars, 1970-1971.
11. S. LANG, *Linear algebra*. Addison Wesley, 1966.
12. J. LELONG FERRAND, *Les notions mathématiques de base dans l'enseignement du second degré*. A. Colin, 1964.

Símbolos utilizados en este libro

$\{a, b, c, \dots, l, m\}$	Conjunto cuyos elementos son los objetos a, b, c, \dots, l, m .
(a_1, a_2, \dots, a_n)	n -plas.
$\mathcal{P}(E)$	Conjunto de las partes del conjunto E .
$\text{card}(E)$	Cardinal del conjunto E .
$A \setminus B$	Diferencia de los conjuntos A y B ; la relación $x \in A \setminus B$ equivale a ($x \in A$ y $x \notin B$).
$f: E \rightarrow F$ $x \mapsto f(x)$	Aplicación de E en F , que envía x sobre $f(x)$.
$\mathcal{F}(E, F)$	Conjunto de las aplicaciones de E en F .
$\binom{n}{p}$ o C_n^p	El número $\frac{n!}{p!(n-p)!}$ (donde $n, p \in \mathbf{N}$, con el convenio que $0! = 1$).
$\sup_{x \in A} x$ o $\sup_E A$	Supremo de la parte A de un conjunto ordenado E .
$\sup(x_1, x_2, \dots, x_n)$	Supremo del conjunto $\{x_1, x_2, \dots, x_n\}$ en un conjunto ordenado E ; en particular: el elemento mayor de este conjunto, si existe.
A^*	Si A representa un anillo, conjunto de los elementos no nulos de A .
\mathbf{N}^*	Conjunto $\mathbf{N} \setminus \{0\}$.
\mathbf{N}_n	Conjunto de los enteros p tales que $0 \leq p \leq n$.
\mathbf{N}_n^*	Conjunto de los enteros p tales que $1 \leq p \leq n$.
\mathbf{Q}_+	Conjunto de los racionales ≥ 0 .
\mathbf{R}_+	Conjunto de los reales ≥ 0 .
\mathbf{Q}_+^*	Conjunto de los racionales > 0 .
\mathbf{R}_+^*	Conjunto de los reales > 0 .
\mathbf{R}_-	Conjunto de los reales ≤ 0 .
E/\mathcal{R}	Conjunto cociente de E por la relación de equivalencia \mathcal{R} .

E/F	Grupo cociente del grupo E por el subgrupo normal F , o: módulo cociente del módulo E por el módulo F .
\mathfrak{S}_n	Grupo de las biyecciones de \mathbf{N}_n^* sobre sí mismo.
\mathcal{A}_n	Grupo de las biyecciones <i>pares</i> de \mathbf{N}_n sobre sí mismo.
$[G : H]$	Índice del subgrupo H de G . (G , grupo).
A/\mathfrak{a}	Anillo cociente del anillo conmutativo unífero A por el ideal \mathfrak{a} .
$\mathbf{Z}/n\mathbf{Z}$	Anillo de las clases módulo n ($n \in \mathbf{N}$).
$a \equiv b \pmod{n}$	Los enteros a y b son congruentes mod n ($n \in \mathbf{N}$).
(a_1, a_2, \dots, a_n)	Ideal engendrado en el anillo conmutativo unífero A por los elementos a_1, a_2, \dots, a_n .
(a)	Ideal principal engendrado por el elemento a (en un anillo conmutativo unífero).
$\text{gr}(A)$	Subgrupo engendrado por la parte A (de un grupo).
$K[X]$	Anillo de los polinomios de una variable, con coeficientes en el anillo (o el cuerpo) K .
$K[X_1, \dots, X_n]$	Anillo de los polinomios con n variables con coeficientes en K .
$\text{grad}(P)$	Grado del polinomio P .
$\text{Vect}(A)$	Subespacio engendrado por la parte A (de un espacio vectorial).
$\mathcal{L}_K(E, F)$	Conjunto de las aplicaciones lineales de E en F (E, F : espacios vectoriales sobre el cuerpo K).
$\text{Aff}(A)$	Subvariedad afín engendrada por la parte A (de un espacio afín).
$\det(u)$ (resp. $\det(M)$)	Determinante del endomorfismo u (resp. de la matriz M).
\tilde{M}	Matriz complementaria de la n -matriz cuadrada M .
${}^t f$	Traspuesta de la aplicación lineal f .
E^*	Dual del espacio vectorial E (sobre un cuerpo conmutativo).
A°	Ortogonal, en E^* , de la parte A del espacio vectorial E (resp. en E , de la parte A de E^*), siendo conmutativo el cuerpo de base.
$\text{GL}(E)$	Grupo lineal del espacio vectorial no nulo E .
$\text{GL}(n, K)$ (o $\text{GL}_n(K)$)	Grupo lineal del K -espacio vectorial K^n ($n \in \mathbf{N}^*$), o, cuando K es conmutativo, grupo de las matrices cuadradas invertibles de orden n sobre K .
$\mathcal{M}_{n,p}(K)$ (o $M_{n,p}(K)$)	Espacio vectorial de las (n, p) -matrices sobre K .
$\mathcal{M}_n(K)$ (o $M_n(K)$)	K -Álgebra de las n matrices cuadradas sobre el cuerpo conmutativo K .
${}^t M$	Traspuesta de la matriz M .
$O(n, Q)$	Grupo ortogonal de la forma cuadrática Q sobre un espacio vectorial de dimensión n .

	Grupo ortogonal de la forma cuadrática $\sum_{i=1}^n x_i^2$ sobre \mathbf{R}^n , o grupo de matrices ortogonales reales de orden n .
$\mathbf{SO}(n, \mathbf{R})$	Subgrupo de $O(n, \mathbf{R})$ formado de los endomorfismos ortogonales de determinante $+1$. O : grupo de las n matrices ortogonales de determinante $+1$.
$\mathbf{U}(n, \mathbf{C})$	Grupo de los endomorfismos unitarios sobre \mathbf{C}^n , para la forma hermitica $\sum_{i=1}^n x_i \bar{y}_i$.
$\mathbf{SU}(n, \mathbf{C})$	O : grupo de las matrices complejas unitarias de orden n . Subgrupo de $\mathbf{U}(n, \mathbf{C})$ formado de los endomorfismos unitarios de determinante $+1$. O : grupo de las matrices complejas unitarias de orden n , de determinante $+1$.
A^\perp	Ortogonal de A (en un espacio vectorial respecto de una forma cuadrática o hermitica).
E_n	Espacio euclídeo de dimensión n .
$\langle x, y \rangle$	Producto escalar de un elemento x del espacio vectorial E y de un elemento y de su dual E^* .
$(x y)$	Producto escalar en un espacio prehilbertiano.
$ x $	Valor absoluto (en \mathbf{R}); módulo (en \mathbf{C}); norma euclídea (en E_n).
(x_1, x_2, \dots, x_n)	Producto mixto de n vectores x_i (en E_n orientado).
$x_1 \wedge x_2 \wedge \dots \wedge x_{n-1}$	Producto vectorial de $n-1$ vectores x_i (en E_n orientado).
M^*	Matriz adjunta de la n -matriz cuadrada compleja M .
φ^*	Endomorfismo adjunto del endomorfismo φ (en un espacio hermitico).
${}^t\varphi$	Endomorfismo traspuesto del endomorfismo φ (en un espacio euclídeo).

Índice alfabético

- Abeliano* (grupo — 58).
Absurdo (demostración por el — 5).
Acotado superiormente (elemento — 33).
Adjunta (matriz — 469).
Adjunto (359).
Adjunto (endomorfismo — 472-478).
Afin (aplicación — 310, — recta — 312, espacio 312, — grupo 311, — plano — 312, — variedad — engendrada 313, subvariedad 312).
Alembert (teorema de — 168).
Álgebra (estructura del — 135, — de polinomios 138-170, — de endomorfismos 262, — de matrices cuadradas 319).
Algoritmo (de Euclides — 149).
Alternada (forma multilineal — 349).
Alternado (grupo — 80, polinomios — 179).
Anillo (91, — conmutativo 92, — de Boole 99, — de los endomorfismos de un espacio vectorial 299, — de los enteros módulo n 98, — de los polinomios 171, — euclídeo 538, — factorial 503, — integro 98, — principal 112, — producto 101, — cociente 105, sub — 100, — unífero 91).
Antisimetría (propiedad de — 30).
Aplicación (15, — semilineal 457, — biyectiva 17, — bilineal 343, — compuesta 16, — creciente 32, — decreciente 32, — idéntica 16, — inyectiva 16, — involutiva 17, — lineal 128-280, — p -lineal simétrica 343, — racional 523, — recíproca 17, — epiyectiva 17).
Argumento (de un número complejo — 123).
Asociados (elementos — en un anillo unífero íntegro y conmutativo — 109).
Asociativa (ley — 53).
Autoadjunto (endomorfismo 472).
Automorfismo (48, — de grupo 59, — internos 60).
Autopolar (sistema de puntos — con relación a una cuádrica 487).
Base (— afin 314, — canónica 296, — de un espacio vectorial 287, — de un módulo 133, — dual 303, — ortogonal 426, — ortonormal 426, — teorema de la — incompleta 294).
Base incompleta (teorema de la — 294).
Bezout (teorema de — 146).
Bidual (302).
Bilineal (aplicación — 343, — forma — 415).
Binomio (fórmula del — 93).
Característica (— de un anillo 107, — de un cuerpo 113, — determinante — 377, polinomio —, subespacio — 384).
Cardan (fórmulas de — 201).
Cardinal (— de un conjunto 42).
Cauchy-Schwarz (desigualdad de — 435-442-461-464).
Cayley (método de — 225, teorema de — Hamilton 396).
Centro (— de un anillo 135, — de un grupo 76).
Cerrado (cuerpo algebraicamente, 168).
Cíclico (— grupo 70).
Ciclos (— ligado a una hipersuperficie 515, descomposición de una permutación en — 86).
Clases (— por la derecha 72, — por la izquierda 72, — de equivalencia 26, — de conjugación

- 85, — de p -transitividad 84, — ecuación de 83, — módulo n , 98).
- Cociente* (conjunto — 26, — en la división euclídea 143, — en la división según las potencias crecientes 257-261).
- Codimensión* (296).
- Coeficientes* (—de una combinación lineal 130, — de una matriz 315).
- Colectivizante* (— relación 10).
- Colineal* (282).
- Columna* (316).
- Combinación* (39, — con repetición, 40).
- Combinación lineal* (130).
- Compatible* relación — 50, sistema — 324).
- Complementario* (—de un conjunto 10, — de una matriz 359).
- Compuesta* (— aplicaciones 16).
- Conjunción* (4).
- Conjugado* (— de un número complejo 120, —subgrupo 82).
- Conjunto* (9, — de llegada 13, — de salida 13, — de definición 13, — enumeración 36, — ordenado 30).
- Conmutativa* (ley, 55).
- Contradictoria* (— teoría 3).
- Coordenadas* (— relativas a una base 133, formas — 303).
- Correspondencia* (13, — algebraica 511, — homográfica 512).
- Cota* (— inferior, — superior 33).
- Cota superior* (33).
- Cramer* (fórmulas de — 373, sistema de — 372).
- Creciente* (— aplicación 27, — aplicación estrictamente 32).
- Cuádrica* (478).
- Cuantificador* (— existencial 7, — universal 7).
- Cuatriones* (576).
- Cuerpos* (113, — de los complejos 120, — de las fracciones 117, de los racionales 116, — de los reales 116, — sub 114).
- Curva* (— algebraica 522, — unicursal 523).
- Decreciente* (— aplicación 32).
- Degenerada* (forma bilineal — 418, forma hermitiana — 458, transformación homográfica — 512).
- Derivada* (— de un polinomio 161, — de una fracción racional 273, — parcial 174).
- Desarrollo* (—de un determinante 355, — en serie formal de una fracción racional 263).
- Descomposición* (— de un polinomio en factores irreducibles 155, — de una forma hermitiana 460, — en cuadrados de una forma cuadrática 427-428-429-432, — en elementos simples de una fracción racional 253).
- Descomposición canónica* (—de una aplicación 28, — de una aplicación lineal 283, — de un homomorfismo 51, — de un homomorfismo de anillo 106, — de un homomorfismo de grupos 75).
- Desigualdad* (— de Schwarz 435-442-461-464, — triangular 435-462).
- Desplazamiento* (499).
- Determinante* (— borde 370, — característico 377, — de un endomorfismo 352, — de una matriz 353, — de una matriz circular 367, — de n vectores 349, — de Vandermonde 364, — menor 357, — principal 374).
- Diagonal* (—de un producto 23, — de una matriz 316).
- Diagonalizable* (endomorfismo — 388, matriz — 388).
- Diagrama* (28, — conmutativo 29).
- Diferencia* (— de dos conjuntos 12).
- Dimensión* (— de un espacio vectorial 293, espacio de —finito 290, teorema de la — 292).
- Dirección* (— de una variedad afín 312).
- Discriminante* (— de un polinomio 237, — de una forma bilineal 418).
- Distinto* (— subgrupo 74).
- Distributiva* (— ley 89).
- Disyunción* (3, 5).
- División* (— euclídea de polinomios 143, según las potencias crecientes 257).
- Divisor de cero* (97).
- Dominio* (— de integridad 97, — de operadores 48).
- Dual* (— de un espacio vectorial 301).
- Ecuación* (—algebraica 197, — de las clases 84).
- Eisenstein* (criterio de — 542).
- Elemento* (—invertible 96, — irreducible 111, — maximal 35, —máximo 32, — neutro 54, — nilpotente 536, — unidad 92).
- Elementos* (— asociados en un anillo 110).
- Elementos simples* (— en la descomposición de una fracción racional 253, — de primera especie 266, — de segunda especie 268).
- Eliminación* (215).
- Endomorfismo* (48, — adjunto 472, — autoad-

- junto 472, — de un grupo 99, — diagonalizable 388, — hermítico 472, — nilpotente 408, — normal 474, — ortogonal 447 — simétrico 477, — ortogonal 447, — simétrico 477, — traspuesto 477, — unitario 466).
- Engendrado* (grupo — 63, ideal — 104, subespacio vectorial — 288, submódulo — 131, variedad afín — 312).
- Equipotentes* (— conjuntos 42).
- Equivalentes* (formas cuadráticas — 427, matrices — 337, relaciones — 4).
- Escalar* (282, producto 434).
- Espacio* (— afín 309, — euclídeo 433, — prehilbertiano 460, — producto 284, — cociente 282, — vectorial 281).
- Estable* (parte — 49, subespacio 394).
- Estructura* (47, — inducida 50, — cociente 50).
- Euclides* (algoritmo de — 149).
- Euclideo* (espacio — 433).
- Euler* (ángulos de 339, — criterio de — 166, función de — 109, fórmulas de — 109-123, identidad de — 176).
- Exponenciales* (122).
- Factores* (— invariantes de una matriz cuadrada 411).
- Factorial* (anillo — 503).
- Factorización* (ver *Descomposición*).
- Familia* (20, — finita, 20, — generadora 131, — intersección y reunión de — 120, — libre 132, — ligada 132, — de conjuntos 23, sub 20).
- Fermat* (teorema «pequeño» de — 117).
- Ferrari* (método de — 206).
- Fielmente* (grupos que operan — 84).
- Fila* (— de una matriz 316).
- Finito* (conjunto — 37).
- Forma* (— bilineal 415 — bilineal simétrica 418, — coordenada 303, — definida positiva 430, — hermítica 455, — hermítica definida, positiva 460, — lineal 301, — p -lineal 343, — p -lineal alternada 345, — multilineal 344, — cuadrática 418, — reducida de una matriz 403, — sesquilineal 455).
- Forma reducida de una matriz* (403-411).
- Formas cuadráticas equivalentes* (427).
- Fracciones* (cuerpos de las — 120, descomposición de las — racionales 253, forma irreducible de las — racionales 250, — racionales 249).
- Función* (15, — coordenada 303, — de Euler 109, — homográfica 512-534, — polinomio 157, — racional 251).
- Funcional* (grafo — 15).
- Gauss* (método de — 374, teorema de — 147-506).
- Generadora* (familia —, parte — 115).
- Generadores* (— de $\mathbb{Z}/a\mathbb{Z}$ 70, sistema de — de un módulo 131).
- Giro* (493).
- Grado* (— de un polinomio 139, — parcial 173, — total 173).
- Grafo* (13, — funcional 15).
- Grupo* (57, — abeliano 58, — afín 311, — alternado 80, — circular 534, — cíclico 70, — de permutación 58-65, — de isotropía 82, — de las unidades 97, — finito 76, — lineal 299, — producto 65, — cociente 66, — ortogonal 448, — simple 533, — especial ortogonal 450, — especial unitario 468, — simétrico 58, — unitario 466).
- Hamilton-Cayley* (teorema de — 396).
- Hermítico* (endomorfismo — 408, forma — a 455, matriz — a 456).
- Hiperplano* (— afín 313, — vectorial 305).
- Hipersuperficie* (— algebraica 515).
- Homogéneo* (polinomio — 175, sistema lineal — 372).
- Homografía* (511).
- Homomorfismo* (48, — de anillo, — de grupos 59, — exponencial 122).
- Ideal* (103, — por la derecha, por la izquierda 536, — bilátero 536, — engendrado 104, — maximales 115, — primo 507, — principal 104, — suma 105).
- Identidad* (— de Euler 175, — formal o funcional de polinomios 160).
- Imagen* (— de un homomorfismo de grupos 62, — de una aplicación 14, — de una aplicación lineal 282, — directa de una parte 18, — recíproca de una parte, 18).
- Implicación* (3).
- Inclusión* (9).
- Incógnitas* (372, — principal 374).
- Indecidable* (relación — 3).
- Independencia lineal* (132).
- Indeterminada* (141-172).
- Índice* (— de un subgrupo 72).

- Inducida* (estructura — 50).
Inercia (ley de — 429-460).
Infinito (axioma del — 13, conjunto — 13).
Intercambio (teorema del — 291).
Intersección (— de dos conjuntos 10, — de una familia de conjuntos 20).
Intransitivamente (82).
Invariantes (ecuaciones a los — 484, factores — de una matriz 411, subgrupos — 72).
Inversa de un elemento (— en un anillo 96, — en un grupo 58).
Inversiones (número de — de una permutación 79).
Invertible (elemento — 96, endomorfismo — 299, matriz — 320).
Involución (17).
Inyectiva (aplicación — 17).
Irreducible (elemento — en un anillo 111, polinomio — 154).
Isometría (489).
Isomorfismo (48, — de grupos 59).
Isotropía (grupo de — 81).
Isótopo (subespacio —, vector — 425).
Jordan (matriz de — 414, reducida de — 414).

Kronecker (símbolos de — 295).

Lagrange (fórmula de interpolación de — 542).
Leibnitz (fórmula de — 161).
Ley de composición (— externa 47, — inducida 49, — interna 47, — cociente 50).
Libre (familia — 132, parte — 133-287, parte afín — 314).
Ligada (familia — 132, parte — 132).
Lineal (aplicación — 128-282, combinación — 130, forma — 301, independencia 132).

Matriz (— adjunto 469, — antisimétrica 325, — asociada a una forma bilineal 416, — asociada a una forma hermítica 455, — de cambio 335, — cuadrada 316, — circular 367, — columna 328, — complementaria 359, — de una aplicación lineal 326, — de un sistema lineal 372, — de Jordan 414, — diagonal 316, — diagonalizable 388, — elemental 317, — equivalente 337, — extraída 370, — hermítica 456, — invertible 320, — nilpotente 408, — ortogonal 446, — producto 317, rango de una 369, — escalar 320, — semejantes 338, — simétrica 325, — triangular 323, — unitaria 469, — unidad 320).
Maximal (elemento — 35, ideal — 116).
M.C.D. (111, — de dos polinomios 145).
MCM (111, — de dos polinomios 148).
Medio (hiperplano — 445).
Menor (— de un coeficiente de una matriz cuadrada 357).
Métrica (— euclídea 435, — hermítica 461).
Minimal (elemento — 35, — polinomio — 396).
Mixto (producto — 453).
Módulo (clases — n 98).
Módulo (127, — de un número complejo 121).
Moivre (fórmula de De — 123).
Monótona (aplicación — 32).
Morfismo (48).
Multilineal (aplicación —, forma — 343).
Multiplicidad (orden de — de una raíz 163, orden de — de un valor propio 389).

Negación (3).
Neutro (elemento — 54).
Newton (fórmula de — 190).
Nilpotente (elemento — 536, endomorfismo — 409).
Norma (— euclídea 436, — hermítica 463).
Normal (endomorfismo — 474).
Normalizado (polinomio — 137).
Normalizador (85).
n-plas (11).
Núcleo (— de un homomorfismo de grupos 62, — de una aplicación lineal 282, — de una forma cuadrática 424).
Numerable (— conjunto 43).

Opuesto (— de un elemento con un grupo abeliano 58).
Órbita (82).
Orden (— de un elemento en un grupo 64, — de un polinomio 140, — de un polinomio simétrico 185, — de una serie formal 260, — de multiplicidad de una raíz 163, — parcial 31, — total 31).
Orientación (452).
Ortogonal (automorfismo — 448, base — 426-459, — de una parte de un subespacio 302-423-459, grupo — 448, matriz — 447, proyección — 445-465, suplementario — 437-463, sistema — 439).
Ortogonalidad (— con relación a una forma hermítica 459, — con relación a una forma cua-

- drática 424).
- Paralelas* (variedades afines — 313).
- Paramétrica* (representación — propia 525).
- Pares* (11).
- Parseval* (fórmula de — 432-465).
- Parte* (— de un conjunto 9, — entera de una fracción racional 256, — generador de un módulo 131, — generatriz de un espacio vectorial 287, — libre 133, — polar relativa a un polo 256, — vacía 10).
- Partición* (27).
- Pascal* (relación de — 40).
- Permutación* (17, — circular 86, grupo de las — s, 58 — impar 80, — par 80).
- Peso* (— de un polinomio simétrico 185).
- p-lineal* (aplicación — 343, forma — 343, forma — alternada 345, forma — simétrica 345).
- Polinomio* (137, — alternado 179, — con n incógnitas 170, — característico 384, — cíclico 179, — ciclotómico 543, — de endomorfismo 394, función — 156-177, — homogéneo 175, — irreducible 154, — minimal 396, — normalizado (o unitario) 137, — simétrico 179-180).
- Polo* (— de una fracción racional 251).
- Potencia* (— del continuo 43, — del numerable 43).
- Primitiva* (— de una fracción racional 274, raíz — de la unidad 70).
- Primitivo* (polinomio — 508).
- Primos* (polinomios — entre sí 164).
- Primos entre sí* (elementos — 437).
- Principal* (anillo — 112, determinante — 374, dirección — 479, ecuación — 374, ideal — 104, incógnita — 374).
- Producto* (— cartesiano de conjuntos 11, — de anillos 101, — de espacios vectoriales 284, — de una familia de conjuntos 24, — de grupos 66, — de matrices 317, — de polinomios 137, — mixto 453, — escalar 434-460, — vectorial 453).
- Prolongación* (— de una aplicación 17).
- Propio* (subespacio — 287, valor — 383, vector — 383).
- Proyección* (24, canónica 27, — ortogonal 443-463, — sobre un subespacio 285).
- Racional* (aplicación — 523, función — 251, fracción — 249).
- Raíz* (— de un número complejo 126, — de un polinomio (o de una ecuación) 158, — primitiva de la unidad 72-126, — simple 164).
- Rango* (— de un sistema de vectores 293, — de un sistema lineal 372, — de una aplicación lineal 297, — de una forma bilineal 418, — de una forma hermitica 458, — de una matriz 369).
- Razón doble* (533).
- Recíproca* (aplicación — 17, ecuación — 209-210).
- Recurrencia* (demostración por — 36).
- Recurrentes* (sucesiones — 404).
- Reducción* (de una matriz 403, — de Jordan 414, — de las formas hermiticas 460, — de las formas cuadráticas 372).
- Regular* (elemento — 55, polinomio — 516).
- Relación* (— binaria 14, — de equivalencia 26, — de orden 30).
- Resolvente* (201).
- Resto* (— en la división euclídea 143, — en la división, según las potencias crecientes 257-261).
- Restricción* (— de una aplicación 16, — de los escalares 286).
- Resultante* (— de dos polinomios 218).
- Reunión* (— de dos conjuntos 11, — de una familia de conjuntos 20).
- Rotaciones* (496-497).
- Rouché-Fontené* (teorema de — 378).
- Sarrus* (regla de — 362).
- Schmidt* (procedimiento de — 440).
- Schwarz* (desigualdad de *Cauchy* — 435-442-461, 464).
- Semilineal* (aplicación — 457).
- Serie formal* (260).
- Sesquilineal* (forma — 455).
- Signatura* (— de una forma hermitica 460, de una forma cuadrática 429 — de una permutación 80).
- Simetría* (443-465).
- Simétrico* (— de un elemento 57, grupo — 58).
- Sistema* (— de ecuaciones lineales 372, — de Cramer 373, — homogénea 374, — ortonormal 439).
- Subanillos* (100).
- Subcuerpo* (114).
- Subfamilia* (20).
- Subgrupo* (61, — conjugado 82, — distinguido

- (o invariante o normal) 74, — engendrado 63).
Submódulo (129, — engendrado 131).
Subyacente (conjunto — 49).
Sucesión (23, sub — 23, — estacionaria 23, — recurrente 404).
Suma (— de ideales 105, — de subespacios vectoriales 285, — de submódulos 129, — directa externa de espacios vectoriales 283).
Suplementaria (— de un subespacio vectorial 286, — ortogonal 437-463).
Suspensión de una matriz (322).
Sylvester (determinante y matriz de — 217, ley de inercia de — 429-460).
Taylor (fórmula de — 162).
Tipo (— de una forma cuadrática 373).
Transformada (— de una ecuación 193).
Transitivamente (grupo que opera — 82).
Transitividad (— de una relación 26).
Trasposición (78).
Traspuesta (— de un endomorfismo de un espacio euclídeo 477, — de una aplicación lineal 302, — de una matriz 317).
Traza (— de un endomorfismo 384 — de una matriz 383).
Triangulación (— de una matriz 385).
Tschirnhaus (transformación de — 231).
Unicursal (curva — 525).
Unidad (elemento — 92, grupo de — 97).
Unitaria (automorfismo — 466, grupo — 466, matriz — 467, polinomio — 137).
Vacío (conjunto — 10).
Valor propio (— de un endomorfismo 383, — de una matriz cuadrada 385).
Valoración (— de un polinomio 140, — de una serie formal 260).
Vandermonde (determinante de — 364).
Variaciones (número de — 38).
Vector (281, — columna, — fila 316, — propio 383).
Vectorial (espacio — 281, producto — 453).
Waring (fórmulas de — 195).
Zorn (8).



ISBN - 84 - 291 - 5065 - X

ISBN - 84 - 291 - 5066 - 8